

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2004, 13:18:42 ; Search time 5232 Seconds
(without alignments)
2925.208 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MRSFGGEHCHAPKKEPVKVG.....ATOLVGEONCIRKIFQGPID 420

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgnt 1/USPTO.spool.p/US10041030/runat.09122004.095517.7675/app.query.fasta_1.583
-DB=EST -QPMT=fastcap -SUFFIX=ret -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10041030.@CGN 1.1.6425.@runat.09122004.095517.7675 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=5
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2151.5	94.0	3387	AK030564 Mus muscu
2	2147	93.8	1186	AY409116 Homo sapi
3	2052.5	89.6	1183	AY409118 Mus muscu
4	1926	84.1	3530	AK045515 Mus muscu
5	1912	83.5	2584	AK045673 Mus muscu
6	1681	73.4	1186	AY409117 Pan trogl
7	1632	71.3	1879	AK044418 Mus muscu
8	1626.5	71.0	1242	AY408790 Pan trogl
9	1622	70.8	2572	CR606152 full-leng

10	1621.5	70.8	1242	AY408789	AY408789 Homo sapi
11	1611.5	70.4	1242	AY408791	AY408791 Mus muscu
12	1563	68.3	2474	AK033815	AK033815 Mus muscu
13	1519	66.3	849	BX388547	BX388547 Mus muscu
14	1473	64.3	3531	AK029586	AK029586 Mus muscu
15	1430.5	62.5	957	CN329888	CN329888 AGENCOURT
16	1391	60.7	869	BX688136	BX688136 Mus muscu
17	1363.5	59.5	856	CF241089	CF241089 AGENCOURT
18	1362	59.5	1117	CK027873	CK027873 AGENCOURT
19	1334.5	58.3	1076	BM800644	BM800644 AGENCOURT
20	1314	57.4	809	BX700555	BX700555 Mus muscu
21	1306.5	57.4	764	BP682660	BP682660 Mus muscu
22	1294	56.5	701	CK637810	CK637810 UT-M-HO-
23	1263.5	55.2	738	BP701801	BP701801 UT-M-HO-
24	1261	55.1	768	CO738704	CO738704 SILEO4C21
25	1258	54.9	784	BX749714	BX749714 BX749714
26	1237	54.0	1229	BC026695	BC026695 Mus muscu
27	1203.5	52.6	752	BP709958	BP709958 Mus muscu
28	1189	51.9	742	BX774759	BX774759 BX774759
29	1185	51.7	1006	BQ049269	BQ049269 AGENCOURT
30	1161	50.7	766	CF539872	CF539872 UT-M-BX-
31	1159	50.6	889	BX327874	BX327874 BX327874
32	1154.5	50.4	795	CK638809	CK638809 UT-M-HO-
33	1133	49.5	964	CD301176	CD301176 AGENCOURT
34	1126	49.2	826	BI183297	BI183297 UNL-P-FN-
35	1122	49.0	2342	AK013388	AK013388 Mus muscu
36	1106	48.3	949	BX369994	BX369994 BX369994
37	1098	47.9	893	AL888857	AL888857 AL888857
38	1089	47.6	893	BX410297	BX410297 BX410297
39	1087	47.5	839	BU934702	BU934702 AGENCOURT
40	1083.5	47.3	659	AL873282	AL873282 AL873282
41	1067.5	46.6	851	AL658565	AL658565 AL658565
42	1065	46.5	876	CF251704	CF251704 hdm003-C0
43	1063	46.4	962	BX341733	BX341733 BX341733
44	1061	46.3	593	BJ034307	BJ034307 BJ034307
45	1052	45.9	600	BG803249	BG803249 0222-86 M

ALIGNMENTS

RESULT 1	AK030564	3387 bp	trna	linear	HTC 03-APR-2004
LOCUS	AK030564				
DEFINITION	Mus musculus adult male pituitary gland cDNA, RIKEN full-length enriched library, clone:530432G3 product:pellino 2, full insert sequence.				
ACCESSION	AK030564	GI:26326560			
VERSION	AK030564.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042155				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,				

TITLE
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
20530913
11076861

REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

TITLE
JOURNAL
AUTHORS
5
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
AUTHORS
6 (bases 1 to 3387)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsumi, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submision
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

TITLE
JOURNAL
COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. .3387
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:5330432G13"
/db_xref="taxon:10090"
/clone="5330432G13"
/sex="male"
/tissue_type="pituitary_gland"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
336. .1595
/note="unnamed protein product; peellino 2
[MGD|MGI:1891445, GB|NM_033602, evidence: BLASTN, 99%, match=864]
putative"
/codon_start=1
/protein_id="BAC27024.1"
/db_xref="GI:26326561"
/translation="MPSGKKEPTPNKKPVKRELVVLYGYALPNGDRGRKSRFA
LYKRYASGVSTVTHMSTVPOASKASISRGHSISYLTSLRQVTVVETTHKXDMFD
OYGRSTESIPDVSTVDTVSGCNEDAQIQTOSTISRPACHTVDRNRPYRARIPLAAGFD
SKNTFLGSKAAKWKPKDGMGDLTTNGLVNHPQGGFPEESQPGWREISVCSGVYT
LRETSAQQRGLVESETVLDGSLIDLCGATILMRWDGLFHAPTQKHIALQEI
NARPOCPGLNLTALFPSINRKEVVEEKPMAYLISGHVGHYSHGHSRDTANREBC

FEATURES
source

CDS

ORIGIN
PMCRVTGPGYVPLMIGCEAGFYVDAGPPTHAFTPCGHVCSKSAKYWSQIPLPHGTAA
HAACFPANTQLVGEQNKILFGQVVD"

Alignment Scores:
Pred. No.: 1,536-205 Length: 3387
Score: 2151.50 Matches: 395
Percent Similarity: 96.67% Conservative: 11
Best Local Similarity: 94.05% Mismatches: 13
Query Match: 93.95% Indels: 1
DB: 3 Gaps: 1

US-10-041-030-4 (1-420) x AK030564 (1-3387)

QY 1 MetPheSerProGlyGlnGlnGlnHisCySaLaProAsnLysGluProValLysTyrGly 20
||||| :|||
Db AGTTTTCCTCCCGGCGANAAAGAACCCANCAACCCCAAGANCCGGTGAATACCGG 395

QY 21 GluLeuValValLeuGlyTyrAsnGlyValLeuProAsnGlyAspArgGlyArgArgLys 40
|||||
Db GAGCTGCTGCTCTCGGGTCAATGCTCTTACTTAATGTCAGAGGGGAGAGGAAA 455

QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
456 AGCAGATTTCCTCTATTAAGCGACCTACCGCAGTGTGTCAACCCAGCACATCCAC 515

QY 61 ValIleSerThrProGlnAlaSerLysValIleSerCyLysGlyGlnHisSerLysSer 80
61 :|||
Db ATGCTCCACACCAAGCGCGCCCAAGCCATCACTCCAGAGGACATACAGCATATCG 575

QY 81 TyrThrLysSerArgArgAsnGlnThrValValValGlyTyrThrHisAspLysAspThrAsp 100
||||| :|||
Db TACAGCTGTGTACGAGACGACGAGCGTACTGTGTGAGTACACACGATTAAGACACAGC 635

QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
636 ATGTTTCAGTGTGGGAGGTCAACAAAGCCCATGTGCTCTGTGTCAACAGACGCGTT 695

QY 121 SerGlySerGlnAsnThrAspGluLysGlnIleThrGlnSerThrIleSerArgPheAla 140
696 TCCGCGGTGCAAGAACAAAGAT--GCCAGATCACACAGACCAATCTTAGTTCGGA 752

QY 141 CysAlaGlyLevalCysAspArgAsnGluProTyrThrAlaArgLysPheAlaIleGlyPhe 160
753 TCCAGAGTCTGTGTGACAGAGAACGACCATACACAGACGATTTTCGGGCAAGATTCC 812

QY 161 AspSerSerLysAsnIlePheLeuGlyGlyLysValAlaLysTyrLysAsnProAspGly 180
813 GATTCCTCCAAAATATCTCTTCTTGAGAGAAAGCAAGCAAAATGGAACCCCTGATGGA 872

QY 181 HisMetLysPheGlyLeuThrThrAsnGlyValLeuValMetHisAspProArgGlyGlyPheThr 200
873 CACATGAGATGACCTACTACCAACGCTCTCGGAGAGCACCCGCAAGAGGCTTCACC 932

QY 201 GluGluSerGlnProGlyValTyrArgGlyLysLeuValCysGlyAspValTyrThrLeu 220
933 GAGGAATCCAGCGCTCGAGAGTCTGAGAGAGATCTGTCTGTGGGAGTATACACCTTG 992

QY 221 ArgGlyThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGlnThrAsnValLeu 240
993 CGAGAGACACAGGTCCGGCCAGACAGAGGAGAAAGCTGTGAGAAAGTGAACCAACTCTCG 1052

QY 241 GlnAspGlySerLeuIleAspLeuGlyValAlaThrLeuLeuTPRArgThrAlaAspGly 260
1053 CAAGAGCGCTCCCTATGACCTGTGTGGGCACTCTCTCTGAGAAACCGCATGAGC 1112

QY 261 LeuPheHisThrProThrGlnLysHisIleGlnAlaLeuAspArgGlnLysIleAsnAlaIle 280
1113 CTTTTCACGCTCCCTACTCAGAGACACTTAAGAGCCCTCCGGACGAGATGATCAACACCC 1172

QY 281 ArgProGlnCysProValGlyLysAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
1173 GCACCCGAGTGCCTCGTGGGCTTAAACCTGTGCTTCCACAGCATCAACCGAAGGAA 1232

Oy	301	ValValgIugUjULvYvgInPrOTralaTYrIeusSerCYvgJYLHlaValHIgLIYYrhIS	320
Db	1223	GTCGTGAGAGAGAACACCCCTGGGCATTACTTGACTCGGCCCATGTGCAAGGCTAACAC	1292
Oy	321	AenTPrgLYHtSargSerAspThrgLunlalaenguaArgLUcySPrometCySaRgtThr	340
Db	1293	AGCTGGGGGCATTCGGAGCGCACCGAACCACAAGAGGAGATGTCTCCATGTGCGAGCAT	1352
Oy	361	ProPProTHrtSalalPaheThrProcCYvgJYLHtSVAlCYvsSerGLvLSerAlaYsr	380
Db	1413	CCCCCAATCAGCGTTTACACCCCTCGGGCAGCGTGCTTTCAGAGAAGTCTGCCAAGTAC	1472
Oy	381	TtpSerGlnIleProLeuProHiasglyThrhSalapheHtsalalACYasProBheCYs	400
Db	1473	TGTTGTGAGATATCCCATCGCCCCACGAAAGCAGCGGTTTTACAGCGCGCTGTCCTGTTCG	1532
Oy	401	AlAthrgInLeuValglYglImancYstlELYsLeuIllepHgInglyProllaasp	420
Db	1533	GCCACGCGAGCTGGTGTGGAAcAGAAcGCAATCAATTGATTTTCCAAGTCCAGTGCAC	1592
RESULT 2			
LOCUS	AY409116	1186 bp	DNA linear GSS 16-DEC-2001
DEFINITION	Homo sapiens PEL12 gene, VIRTUAL TRANSCRIPT, partial sequence,		
ACCESSION	AY409116		
VERSION	AY409116.1	GI:39765084	
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1186) Clark,A.G., Gianowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J., Adams,M.D. and Cargill,M. Interfing nonneutral evolution from human-chimp-mouse orthologous gene trios		
TITLE	Science 302 (5652), 1960-1963 (2003)		
JOURNAL	14671302		
PUBMED	2 (bases 1 to 1186)		
AUTHORS	Clark,A.G., Gianowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J., Adams,M.D. and Cargill,M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment.		
COMMENT	Location/Qualifiers		
FEATURES	source		
	.1..1186		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	<.1..>1186		
	/gene="PEL12"		
	/locus_tag="HCM3450"		
ORIGIN			
Alignment Scores:			
Pred. No.:	8.77e-206	Length:	1186
Score:	2147.00	Matches:	394
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	93.76%	Indels:	0
DB:	9	Gaps:	0

US-10-041-030-4 (1-420) x AY409116 (1-1186)

QY	27	TyrAenGlyValAlaLeuProAenGlyAAspArgGlyArgArgLysSerArgPheAlaLeuTyr	46
Db	2	TACATAGGTGCTTTTACCCAAATGAGATAGAGAGCGAGAGAAATAGATTGGCCCTCAC	61
QY	47	LysArgProLysValAaenGlyValAlaLysProSerThrValHisValIleSerThrProGln	66
Db	62	AAAGGCGCCCAAGGCAATGTTGTCAAAACCGACCGTCCATGTGATTCACGCCCCAG	121
QY	67	AlaSerLysAlaIleSerCysValGlyGlnHisSerIleSerTyrThrLeuSerArgAsn	86
Db	122	GCATTCAAAGCTTACAGCTCGCAAGGTCACACAGTATCTCTACACTTGTCAAGGAT	181
QY	87	GlnThrValValAlaGlyTyrThrHisAspLysAspThrAspMetPheGlnValGlyArg	106
Db	182	CAGACTGGGGGGAGGTACACACATGATAGATACGATATGTTTCAGTGGCGAGA	241
QY	107	SerThrGlnSerProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThr	126
Db	242	TCAACAGAAAGCCCTTATCGACTTGTGTGTGCACACACGATTTGTGGCAGCCAAACAG	301
QY	127	AspGluAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAsp	146
Db	302	GACGAAAGCCAGATCACACAGAGCACCATATCCAGTTCCGCTCAGAGATCGTGTGCGAC	361
QY	147	ArgAenGluProTyrThrAlaArgIlePheAlaAlaGlyPheAspSerSerLysAsnIle	166
Db	362	AGGATAGAACTTTCACAGCAGGATATTCGCCCGGATTTGACTCTTCCAAAACACTA	421
QY	167	PheLeuGlyGluLysAlaAlaLysTrpLysAsnProAspGlyHisMetAspGlyLeuThr	186
Db	422	TTTCTTGAGAAAAGCGACAAATGTGAAAAACCCAGCGGCACATGATGGGCTCACT	481
QY	187	ThrAenGlyValLeuValMetHisProArgGlyGlyPheThrGlnLysSerGlnProGly	206
Db	482	ACTAATGAGCGTCTGTGATGATCATCCACGAGGGGGCTTACCGAGAGATCCCAACCCGCG	541
QY	207	ValTrpArgGluIleSerValCysGlyAspValTyrThrLeuAenGluThrArgSerAla	226
Db	542	GTCGCGCCGAGATCTGTCTGTGGAGATGTACACTTGCAGAGAAACCAAGTCCGCC	601
QY	227	GlnGlnArgGlyLysLeuValGlnSerGluThrAsnValLeuGlnAspGlySerLeuIle	246
Db	602	CAGCAACGAGAAAGCTGTGTGAAAGTGAACCAACGCTCGCAGAGACGGCTCCTCATT	661
QY	247	AspLeuCysGlyAlaThrLeuLeuTrpArgThrAlaAspGlyLeuPheHisThrProThr	266
Db	662	GACCTGTGTGGGCGCACTCTCTCTGGAGAACACAGATGGGCTTTTCAATCTCAACT	721
QY	267	GlnLysHisIleGlyValAlaLeuArgGlnGluIleAsnAlaAlaArgProGlnCysProVal	286
Db	722	CAGAAAGCAATAGAGCCCTTCGGCAGAGATTAAACCGCCCGCCCTCAAGTGTCTGTG	781
QY	287	GlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValValGlnGlyLysGln	306
Db	782	GGGCTCAACACCTCGGCTTCCCGACGATCAACAGAAAGAGTGTGTGAGGAAGACAG	841
QY	307	ProTrpAlaTyrLeuSerCysGlyHisValHisGlyTyrHisAsnTrpGlyHisArgSer	326
Db	842	CCCTGGGCAATCTCAGTTGTGGCCAGCTGCACGGGTACCAACTGGGGGCATCGAGT	901
QY	327	AspThrGluAlaAenGluArgGluCysProMetCysArgThrValGlyProTyrValPro	346
Db	902	GACACGAGGCGCAACGAGGAGGTGTCCATGTGCAGAGACTGTGGGCCCCATATGTGCT	961
QY	347	LeuTrpLeuGlyCysGluAlaGlyPheTyrValAspAlaGlyProProThrHisAlaPhe	366
Db	962	CTCTGGCTGGCTGTGAGGACGATTTTATGTACGCGAGACCGGCAACTCATGCTTTC	1021
QY	367	ThrProCysGlyHisValCysSerGluLysSerAlaLysTyrTrpSerGlnIleProLeu	386

Db	1022	ACTCCCTGTGAGACACGTGTGCTCGGAGAAGTGTGCAGAAATTA	TGTCACAGTCCCCTTG	1083
Oy	387	ProHsiGlyThrHisAlaPheHisAlaIaAcysProPheCysAlaThr-GlnLeuValGly		406
Db	1082	CCTCATGTGAACATCAATGATTTTACCCTCTTGCCCTTTCTGGCTAACACAGCTGTGGG		1141
Oy	407	GluGlnAenCysGlyLeuLeuIlePheGlnGlyProIleAsp		420
Db	1142	GAGCAAACCTGCATCAATAATTATTTCCAAAGGTC	CAATTGAC	1183
RESULT 3				
LOCUS	Ay409118		1183 bp	DNA linear GSS 16-DEC-2003
DEFINITION	Mus musculus PELI2 gene, VIRUTAL TRANSCRIPT, partial sequence,			
ACCESSION	Ay409118			
VERSION	Ay409118.1			
KEYWORDS	GSS.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1183) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Interferring nonneutral evolution from human-chimp-mouse orthologous gene trios			
JOURNAL	Science 302 (5652), 1960-1963 (2003)			
PUBMED	14671302			
AUTHORS	2 (bases 1 to 1183) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment.			
FEATURES	Location/Qualifiers			
source	1..1183 /organism="Mus musculus" /mol_type="Genomic DNA" /db_xref="taxon:10090"			
gene	<1..>1183 /gene="PELI2" /locus_tag="HCM3450"			
ORIGIN				
Alignment Scores:				
Pred. No.:	3.07e-196	Length:	1183	
Score:	2052.50	Matches:	375	
Percent Similarity:	97.72%	Conservative:	107	
Best Local Similarity:	95.18%	Mismatches:	8	
Query Match:	89.63%	Indels:	1	
Dbl	9	Gaps:	1	
US-10-041-030-4 (1-420) x Ay409118 (1-1183)				
Oy	27	TyrAnngIyaIaleuProAnngIYAAPArGlyARAgLySeSerArpheAlaLeuTyR		46
Db	2	TACATGTTGCTTATCTTAATGCTGTACAAGGGGACAGAGAAAGCAGATTGCCCTCAT		61
Oy	47	LysArgProLysAlaAenGlyValLysProSerThrValHisValIILSerThProGln		66
Db	62	AAGGGAGCTTAGCCAGCGATGCTCAAAACCAAGCAAAATCCATGATCTCCACACACAG		121
Oy	67	AlaserIyAlalaserCyelnyGlyGlnHisserIleserTyRThrLeuSerArgan		86
Db	122	GGCTTCAGGCCATCATGCTCCAGAGACATCACACATATGTAACGTTGTCAGGAGC		181

LOCUS	AK045515	3530 bp	mRNA	linear	HTC 03-APR-2004
DEFINITION	Mus musculus adult male corpora quadrigenima cDNA, RIKEN full-length enriched library, clone:B2J0208C14 product:ipellino 2,				
Qy	87	GlnthrtAlValValAGlurTrrthHLeaSpLyAspRhrthRbmetPheGlnValGlyArg	106		
Db	182	CAGACGGTAAGTGGTGGAGTACACACACAGATPAAGACACGACATGTTCCAGTGGGACGG	241		
Qy	107	SerThrgLusSerProLLeaPpheValValThrAspThrLLeSerGlySerGlnAspThr	126		
Db	242	TCAACGAAAGGCCCATTTGACTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	301		
Qy	127	AspGlnUaGlnLLeThrgLusSerThrLLeSerArgPheLaCyAsArgLLeValCyAsp	146		
Db	302	GAT---GCCAGATCATCACAGACAGACCATCTCTAGTGTTCGATTCAGATTCGTGTGTAC	358		
Qy	147	ArgAspGlnUaProTrrthrrAlaArgLLePheLaLaLaGlyPheAspSerArgAspLLe	166		
Db	359	AGGACAGAGCCATATACAGACAGCATATATGCGGACGATTCGATTTCCAAAATAATC	418		
Qy	167	PheLeuGlyGlnUaValaLaLaLysTrrLysAspProAspGlyHisMetAspGlyLeuThr	186		
Db	419	TTTCTTGGAGAGAAAGACGACAAATGAAATACCTCTATGACACATGATGACTACT	478		
Qy	187	ThrAsnGlyValLeuValMetHisProArgGlyGlyPheThrgLusSerGlnProGly	206		
Db	479	ACCAATGCTGTCTTAATGATGACACCCGACAGAGAGGCTTACACGAGAAATCCGAGCTCGA	538		
Qy	207	ValTrrPArgGlnLLeSerValCyAsGlyAspValTrrThrLeuArgGlnUaTrArgSerLa	226		
Db	539	GTCGAGAGAAAGATCTGTCTGTGGGAGTGTATACCTTGCAGAGACCAAGTCCGACC	598		
Qy	227	GlnGlnArgGlyLysLeuValGlnSerGlnUaThrAsnValLeuGlnAspGlySerLeuLLe	246		
Db	599	CAGCAGAGGGGAAAGCTGTGTGAAAGTGAACCAAGCTCTGCAGAGCGCTCTCTATT	658		
Qy	247	AspLeuCyAsGlyAlaThrLeuLeuTrrArgThrLaAspGlyLeuPheHisThrProThr	266		
Db	659	GACCTGTGTGGGGCACTCTCTCTGTGAGAACCGCAGATGAGCTTTTCAAGCTCTACT	718		
Qy	267	GlnLysHisLLeGlnLaLeuArgGlnGlnUaLLeAsnLaLaLaArgProGlnLysProVal	286		
Db	719	CAGAAAGACATAGAAAGCCCTCCGACAGAAATCAATCAACCCGACCCAGTGGCCCTGTG	778		
Qy	287	GlyLeuAsnThrLeuLaLLePheProSerLLeAsnArgLysGlnValValGlnLysGln	306		
Db	779	GGCTTAAACCTCTGGGCTTCCCAAGCATCAACCGAAGAAAGTGTGGAAGAAACAG	838		
Qy	307	ProTrrPAlaTrrLeuSerCyAsGlyHisValHisGlyTrrHisAsnTrrGlyHisArgSer	326		
Db	839	CCCGGGGCATACCTGAGCTCGGGCATGTGACAGGTCACACAGAGCTGGGGCATCGAGC	898		
Qy	327	AspThrgLuaLaAsnGlnUaArgGlyLysProMetCyAsArgThrValGlyProTrrValPro	346		
Db	899	GACCCGGAACCAACGAGAGGAGGTGTCCCATGTGCGAGACTGTGGGCCCTCTAGTCTCT	958		
Qy	347	LeuTrrLeuGlyCyAsGlnUaLLeGlyPheTrrValAspLaLLeGlyProProThrHisLaPhe	366		
Db	959	CTCTGGCTGTGGCTGTGGAGGAGGATTTTAATGTGATGCGGACCCCAACTACAGCTTTC	1018		
Qy	367	ThrProCyAsGlyHisValCyAsSerGlnUaSerAlaLysTrrTrrSerGlnLLeProLeu	386		
Db	1019	ACCCCTGGGGGACCGTCTGTTCAGAAATCTGCGCAAGTACTGTGTGACATCCACTG	1078		
Qy	387	ProHisGlyThrHisLaLLePheHisLaLaLLeCyAsProPheCyAsAlaThrgLLeuValGly	406		
Db	1079	CCCCACGGAAACGACGCGGTTCATGCGGCTGTCTCGCTTGTGCGGCACGACGTGTGTGT	1138		
Qy	407	GlnGlnAsnCyValLLeLysLeuLLePheGlnGlyProLLeAsp	420		
Db	1139	GAAACAGAACTGCATCAATGATTTTCCAAGTCTCACTGGAC	1180		

QY 201 GUGUGSerGlnProGlyValTTPArGluIleSerValCYseGlyAsePValTYrThrIleu 220
 DB 759 GAGGATTCACAGCTGAGTCTGAGAGAGATTCCTGCTGAGGATGATACACTTG 818
 QY 221 ArGluThrArgSerAlaGlnGlnArgGlyLYseValGluSerGluThrAenValIleu 240
 DB 819 CCGAGAACAGGTCGCGCCAGCAGAGGGGAAAGCTGTGGAAAGTAGACCAACGTCCTG 878
 QY 241 GlnAseGlySerIleuIleAseIleuCYseGlyAlaThIleuIleuThrArgThrAlaAseGly 260
 DB 879 CAGAGAGGCTCCCTCACTTACCTGTGTGGGCCACCTCTCTGAGGAATTCGAGATGGC 938
 QY 261 LeuPheGlyThrProThrGlnLYseHISleGluAlaIleuArgGlnIleAenAlaIle 280
 DB 939 CTTTTCAGGCTCTACTCAGAAAGCAGTGAAGCCCTCGCAGAGATCAATGACGCC 998
 QY 281 ArgProGlnCYseProValGlyLeuAenThrLeuAlaPheProSerIleAenArgLYseGlu 300
 DB 999 CAGACCCAGTGGCCCGTGGGCTTAAACCCCTGCTCCCGCATCAACCGAAGAA 1058
 QY 301 ValValGluGluLYseGlnProThrAlaTYrLeuSerCYseGlyHISValHISGlyTYrHIS 320
 DB 1059 GTGTGTGAAGAAGAGAGCCCTGGGAGTACCTGAGCTGGCCATGTGACGGCTACAC 1118
 QY 321 AenThrGlyHISArgSerAsePThrGluAlaAenGluArgGluCYsePrometCYseArgThr 340
 DB 1119 AGCTGGGGCCATCGAGCCAGCAGAGCAAGCAAGAGAGAGTCTCCATGTGCAGACT 1178
 QY 341 ValGlyProTYrValProIleuThrLeuGlyCYseGluAlaGlyPheTYrValAsePAlaGly 360
 DB 1179 GTGGGGCCCTACGCTCTCTGTGGCTGTGAGGCGAGATTTATGTGCATGGCGGA 1238
 QY 361 ProProThrHISAlaPheThrProCYseGlyHISValCYseSerGluLYseSerAlaLYseTYr 380
 DB 1239 CCCCACACTCAGCCTTTCACCCCTGCGGCGACGTCGTTCAGAGAAGTCGCAAGTAC 1298
 QY 381 TPSeSerGlnIleProIleuProHISGlyThrHISAlaPheHISAlaIleCYseProPheCYs 400
 DB 1299 TGTGTCCAGATCCCATCTGCCCGCCAGAGCAGCGCTTTCAGCCCGCTGTCCGTTCTGC 1358
 QY 401 AlaThrGlnLeuValGlyGluGlnAenCYsIleLYseIleuIlePheGlnGlyProIleAse 420
 DB 1359 GCCACCGACGTGGTGTGTGAACAGACTGATCAATGATTTTCCAAAGTCCAGTGCAGC 1418
 RESULT 5
 LOCUS AK045673 2584 bp mRNA linear HTC 03-APR-2004
 DEFINITION Mus musculus adult male corpora quadrigemina cDNA, RIKEN
 full-length enriched library, clone:B230220E21 product:pellino 1,
 full insert sequence.
 ACCESSION AK045673
 VERSION AK045673.1 GI:26337532
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, Y.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multichipillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 PANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL NATURE 409, 685-690 (2001)
 REFERENCE 5
 TITLE The PANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 JOURNAL ANALYSIS OF BIOLOGICAL DATA 1, 1-10 (2002)
 REFERENCE 6
 TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 NATURE 420, 563-573 (2002)
 JOURNAL MEDLINE 12025841
 PUBMED 12025841
 REFERENCE 7
 TITLE Direct Submission
 JOURNAL SUBMITTED (16-JUN-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp,
 URL: http://genome-gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 COMMENT cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome-gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/
 URL: http://fantom.qualifiers
 1. 2584
 location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM.DB:B230220E21"
 /db_xref="taxon:10090"
 /clone="B230220E21"
 /sex="male"
 /tissue_type="corpora quadrigemina"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 426. 1682
 /note="unnamed protein product; pellino 1
 (MGI:1914495, GB|NM_023324, evidence: BLASTN, 99%,
 match=2228)
 putative"
 /codon_start=1
 /protein_id="BAC32452.1"
 /db_xref="GI:26337533"
 /translation="MFSPDQENHPSPKAPVKYGEILVLGNGSLPNGDRGRKRSFALP

KPRKANGKPSVTHIACTPOAAKAIKSNKODHSISYTLSPAQTVVETVTHDSNTDMPQI
 GRTSEPIDFVDTIVPGOSNSDTSQVOSTISRFACKIICESPSPPTATITAGDS
 SKNIPIGEKAKAKWTSDQMDGLTNGVLMFRNGTEBSKGIWREISVCNVSLS
 RETRBAQORKNWEIETNLODGLSIDLCATILMLTAELSHTPVTKHLEJROBIN
 AAPQCPVEFNTLAPPSMKRKYDVDEKQPVYVNCGHVYHNMKNKEEDGDRCP
 MCRSPGVPLMTGCEAGFYVDAGPPTPHASPCGHVCSKEXTAVMSQIPLPHOTHTFH
 AACPFCAHQALAGSGYIRLIFQGRPLD"

ORIGIN

Alignment Scores:

Pred. No.:	1,58e-181	Length:	2584
Score:	1912.00	Matches:	341
Percent Similarity:	90.48%	Conservative:	39
Best Local Similarity:	81.19%	Mismatches:	38
Query Match:	83.49%	Indels:	2
DB:	3	Gaps:	1

US-10-041-030-4 (1-420) x AK045673 (1-2584)

QY 1 MetPheSerProGlyngluGluHisCysAlaProAsnlysgluProVallystyGly 20
 |||||
 DB 426 ATGTTCCTCCATCAGAAAATCAT-----CCTTCCAAAGCCCACTAAATATGGC 479
 QY 21 GluLeuValAlaLeuGlyTyraanglyAlaLeuProAsnlyAspArglyArgly 40
 |||||
 DB 480 GAACTCTGTCTTGGATATTAATGATCTCTCCCAACGGTGTATGAGAGAGAGAAA 539
 QY 41 SerArgPheAlaLeuTyryllyAsArgProlyAlaAsnlyVallyAsProSerThrValHis 60
 |||||
 DB 540 AGTAGAGTTGCTTTGTTTAAAGCCTTAAGCAAAATGGGTGAAGCCTTACACCGTGAC 599
 QY 61 ValIleSerThrProGlnAlaSerlyAlaIleSerCysAllysglyGlnHisSerIleSer 80
 |||||
 DB 600 ATTCAGATGATCTCTCAGGCTGCCAAGCAATAGCAACAGAACCAAGCATACATATCA 659
 QY 81 TyThrLeuSerArgAsnGlnThrValValGluTyThrHisAspLyAspThrAsp 100
 |||||
 DB 660 TATACCTTATCTGAGCCCAAGCGGTGTGTGAATATCTAGTACAGACACATGAT 719
 QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
 |||||
 DB 720 ATGTTTCAGATGTGGTGGTCACTGAAGATCTTATGATTTTGTATGATGACACCGCT 779
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 |||||
 DB 780 CCTGGAAGTCAGATATTCGACAGCAGCTCACTCAAAACCACTATATCAAGATTGGCC 839
 QY 141 CysArgIleValCysAspArgAsnGluProTyThrAlaArgIlePheAlaAlaGlyPhe 160
 |||||
 DB 840 TGTAGATCATATGTAGCGCAGACCCCTTACAGCTCGATTAATGCTGACGGGTTT 899
 QY 161 AspSerSerLyAsnIlePheLeuGlyGluLeuAlaAlaIleTyTrpLyAsnProAspGly 180
 |||||
 DB 900 GATTATTCAAAACATCTTCTTGGGAGAGAGGCTGCCAAGTGGAAAGACATCGATGGG 959
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMethIAspProArgGlyGlyPheThr 200
 |||||
 DB 960 CAGATGATGGCTTGACCACTAAATGAGGTTCTGTGATGATCACACTAATGGGTTTACA 1019
 QY 201 GluGluSerGlnProGlyValITPArgGluIleSerValLysGlyAspValTyThrIleu 220
 |||||
 DB 1020 GAAAGACTCCAAACCTGGAATATGAGAAATATCATATGTGGAACGCTTCAGCTG 1079
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLyLeuValGluSerGlnThrAsnValLeu 240
 |||||
 DB 1080 CGTGAACCAAGATCAGCTCAGAGAGAGAGAGAGTGTGAAATTGAAACCAATCAGCTA 1139
 QY 241 GlnAspGlySerLeuIleAspLeuCyrglyAlaThrIleuLeuTrpArgThrAlaAspGly 260
 |||||
 DB 1140 CAAGATGGCTCTTAATATGACCTTGTGTGCAACCTTGTGTGCGCTTACGCAAGAGCG 1199
 QY 261 LeuPheHisThrProThrGlnLyHisIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
 |||||

DB 1200 CTTTCCCATATCTCTATCTGTGAAGCATTAAGAACTTTAAGACAGAGATCAATGCACT 1259
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLyGlu 300
 |||||
 DB 1260 CGGCCCCAGAGCCCTCTAGGGTTTCAACACACTGACCTTCCCGCAGTATGAAGAGAGAGAT 1319
 QY 301 ValValGluGluGluLeuProThrAlaTyThrLeuSerCysGlyHisValHisGlyThrHis 320
 |||||
 DB 1320 GTTGTAGTGAAGAAAGAACCAATGAGTATTAATCAATGCGGCGCATGCTCATGGTATCAT 1379
 QY 321 AenTrpGlyHisIleArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
 |||||
 DB 1380 AACTGGGGAACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1439
 QY 341 ValGlyProTyryllyAlaProLeuThrLeuGlyCysGlyAlaGlyPheTyryllyAlaAla 360
 |||||
 DB 1440 GTTGTGCTCATATGCTCTGT 1499
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlnLysSerAlaTyryl 380
 |||||
 DB 1500 CTTCCCAACCCATGCTTGTAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1559
 QY 381 TrpSerGlnIleProLeuProHisGlyTyThrHisAlaPheHisAlaAlaCysProPheCys 400
 |||||
 DB 1560 TGGTCCCAATGCCACTTCTCTATGATGACGACACTTTTCATGACACCTGCCCTTCTGT 1619
 QY 401 AlaThrGlnLeuValGlyGluGlnAsnCyryllyLeuIlePheGlnGlyProIleAsp 420
 |||||
 DB 1620 GCATCATGATGTGGCTGTGTGAACAAAGCTATATCATGACTTATTTTCAAGAGACCTTAAGAC 1679
 RESULT 6
 AY409117 1186 bp DNA linear GSS 16-DEC-2003
 LOCUS AY409117
 DEFINITION Pan troglodytes PEL12 gene, VIRUTAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY409117
 VERSION AY409117.1 GI:39765085
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
 REFERENCE 1 (bases 1 to 1186)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,
 Todd,M.A., Tannenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 JOURNAL
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1186)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,
 Todd,M.A., Tannenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submision
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 JOURNAL
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1..1186
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>1186
 /gene="PEL12"
 /locus_tag="HCM3450"
 ORIGIN
 Alignment Scores: 1.04e-158 Length: 1186

Score:	1661.00	Matches:	317
Percent Similarity:	80.46%	Conservative:	0
Best Local Similarity:	80.46%	Mismatches:	77
Query Match:	73.41%	Indels:	0

US-10-041-030-4 (1-420) X AY409117 (1-1186)

Oy	2	TYrAsngIyAlaLeuProAengIyAspARvgIyATsGATyGySseRATrgPheAlaLeuTYr	46
Db	2	TACATGTCCTTTACCCAAATGAGATAAAGACGAGAGAAAGTAAATTTGGCCCTTAC	61
Oy	47	LYsArGrProIySaIaAsngIyValIySProSerThValIhIeValIleSerThProGln	66
Db	62	AAGGGCCCAAGGCAATAGGTGTCAAAACCAACACTGTCAATGTATATCCACCCCCAG	121
Oy	67	AlAserIySaAlIleSerCybIySgIyGlnIhIseRtIleSerTYThLeuSerATrgsn	86
Db	122	GCGCCAAAGCTTCACGTGCAAAAGGTCAACACAGTATCTTCAACTTGTCAAGGAAT	181
Oy	87	GlnThrValIValIaGluTYThrThIAsApIyAspArThAspMeCPheGlnValIyArg	106
Db	182	CAGACTGTGGTGGAGTACACACATGTATAAAGACACGAAATTTTTCAGANNNNNNNN	241
Oy	107	SeRThIGuSerProIleAspPheValIThAspThrIleSerGlySerGlnAsnThr	126
Db	242	NN	301
Oy	127	AspGluAlaGlnIleThrGlnSerThrIleSerATrgPheAlaCybATrgIleValCyAsp	146
Db	302	NN	361
Oy	147	ArgAsngIuProTYThrAlaATrgIlePheAlaIaGlyPheAspSerSerIyAsnIle	166
Db	362	NN	421
Oy	167	PheLeuGlyGluIySaIaAlaIySTrIySaenProAspGlyhIhIeMeCAspGlyLeuThr	186
Db	422	NNNNNNNNNAAGGAGAGCAAGTGAABAAACCCGACGGCACTAGATGGGCTCACT	481
Oy	187	ThrAsngIyAlaLeuValIeChIAsProATrgIyGlyPheThrGluGluSerGlnProGly	206
Db	482	ACTAATGGCCTTCGTGGATGCATCCACGAGGGGCTTCAACGAGAGTCCACCCGGG	541
Oy	207	ValITrpArgrGluIleSerValCybGlyAspPValTYThrThIeATrgIuThATrgSerAla	222
Db	542	GTCCTGGGCGAAGATCTGTCTGTGGAAGATGTATACCTTGCAGAAACAGAGTCGGCC	601
Oy	227	GlnGlnArGlyIyLySleuValGluSerGluThrAsnValIeGlnAspGlySerIleIle	246
Db	602	CAGCAACGAGAAAGCTGGTGAAGATGAGCAACACTCTGCAGAGAGGGTCCCTCTTT	661
Oy	247	AspLeuCybGlyAlaThrIleuLeuTYThATrgThrAlaAspGlyLeuPhehIhThProThr	266
Db	662	GACCTGTGTGGGGCACTCTCCCTCGAAMACAGACAGTGGGCTTTTTCATATCCAACT	721
Oy	267	GlnIyAsIleIleGluIaIeATrgGlnGluIleAsnAlaIATrgProGlnCybProVal	286
Db	722	CAGAGACATTAAGACCTTCGGGACGAGATTAAGCCGGCCGGCTCAAGTGTCTGTG	781
Oy	287	GlyLeuAsnThrIleuAlaPheProSerIleAsnATrgLySgluValIaGluGluLyGln	306
Db	782	GGGCTCAACACCTGGCTTCCCAACATCAACAGGAAGAGTGTGTGAAGAGAAACAG	841
Oy	307	ProTTrpAlaIyTrIeuserCybGlyhIhIeValIhIeGlyTYThIAsenTTrGlyhIAsTrSer	326
Db	842	CCCTGGGCAATATCTCAAGTTGTGGCCACGTGCACGGGTATCCCAACTGGGGCCATCGAGT	901
Oy	327	AspThrGluAlaAsngIuATrgGluCybSPrometCybATrgThValGlyProTYValPro	346
Db	902	GACACGGAGGCCAACGAGAGAGTGTCCANNNNNNNNNNNNNNNNNNNNNNNNNNNN	961
Oy	347	LeuTrIreuGlyCybGluAlaGlyPheTYrValAspAlaGlyProToThrThIAlaPhe	366

Db	962	CTCtGGCTGCGTGTGGAGGATTTTAATGAACAGACGCCAACCACNNNGCTTC	1021
Oy	367	ThPrProcysglYhiSvAlCySeSerGlutylSeSerAlaYstTyTrPSeSerGlnIleProLeu	386
Db	1022	ACTCCCTGTGGACACGCTGTCTCGGAAAGTCTGCAAAATACTGCTTCAGATCCGCTTG	1081
Oy	387	ProHISgLYThrhiSaIaphneHiSaIaAlaCYseProPhcCYbaLaThrGInLeuValGIly	406
Db	1082	CCTCATGGAACtCAcTGcATTTCACGGCTGCTTGCCTTCCTGTGTGTATACAAGCTGGTGGG	1141
Oy	407	GluGlnAnSpCyStIelySeIeuIllepheGlnGIlyProIleLap	420
Db	1142	GAGCAAAACtGCATCAATAATTATTTTCAGAAGTCCAATTGAC	1183
RESULT 7 AKO44418 LOCUS DEFINITION	AKO44418	1879 bp mRNA linear HTC-03-APR-2004 Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930011IL7 product:similar to PELLINO 2 (PELLINO [DROSOPHILA] HOMOLOG 2) (Homo sapiens), full insert sequence.	
ACCESSION VERSION KEYWORDS SOURCE	AKO44418 AKO44418.1 GI:26336455		
ORGANISM	HTC; CAP trapper. Mus musculus (house mouse)		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL MEDLINE PUBMED	Meth. Enzymol. 303, 19-44 (1999) 99279253 10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL MEDLINE PUBMED	Genome Res. 10 (10), 1617-1630 (2000) 20499374 11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Niishi, K., Kitsumai, T., Tashiro, H., Itoh, A., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishire, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaichiagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunuma, S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL MEDLINE PUBMED	Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861		
REFERENCE	4		
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL REFERENCE	Nature 409, 685-690 (2001)		
AUTHORS	5		
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
JOURNAL REFERENCE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)		
AUTHORS	6 (bases 1 to 1879) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hoti, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, U., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kuwahara, C., Matsuyama, T., Miyazaki, A., Murata, M.,		

TITLE
JOURNAL

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Kahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

COMMENT

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Retina RNA was provided by Dr. Stefano Gustinich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

FEATURES

source

CDS

```

1..1879
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A930011L17"
/db_xref="taxon:10090"
/clone="A930011L17"
/tissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
104..1441
/note="unnamed protein product; putative
similar to PELLINO 2 (PELLINO (DROSOPHILA) HOMOLOG 2)
(Homo sapiens) (SPTR|09HAT8, evidence: FASTV, 71.1%ID,
99.2%length, match=1245)"
/codon_start=1
/protein_id="BAC31913.1"
/db_xref="GI:26336456"
/translation="MVLGNPDVGSFRTSDLOHPSQSGSCILSPGEALAGEPIKY
GELIVGVNGLASGDKGRRLSRPHANGVPRDWHNSTLVKXASNRQCH
SISFTLSHVSIVETHTDSKDMFOIGSTEMIDFVDTDFSGGATEGSAOSTI
SRACILCDRRPPTARIYAAGFDASSNIFGERAKKRTDPGLMDGLTNGVLMH
PAGGSEDSAPKWRREISVCGNYTTIRDSRSNRQKRYLVENENVTQDSLDLCAT
LMKRTFAGLRATFLKOLEKORSEANARPOCFVGLSTLAFSPAPKRTAPRQDQW
YVRCGHVHGHWGCRREQPORECRPLRLVGPVPLMGDAAGCLDGPSPSHFA
PCGHVSEKTRAYMAQTPRLPHGTNAFHAAFCFGAWLTGELGCVRLIFQSPILD"

```

ORIGIN

Alignment Scores:

Pre. No.: 1,876-153 Length: 1879
 Score: 1632.00 Matches: 297
 Percent Similarity: 81.58% Conservative: 44
 Best Local Similarity: 71.05% Mismatches: 73
 Query Match: 71.27% Indels: 4
 DB: 3 Gaps: 3

US-10-041-030-4 (1-420) x AK044418 (1-1879)

QY 4 ProG1yG1ng1uG1uh1sCy5a1aProaenlyeG1uProValyETyG1yG1uleVal 23
 DB 194 CCGGGGGAAGAA-----GCACGGCAGCGGAGAGACCCATCAAGTATGTTGAATCATC 247
 QY 24 ValLeuG1yTyRaenG1yLaLeuProaenG1yAaPaRgG1yAaRg1yAaRg1ySeraRgPhe 43
 DB 248 GTTCTGGGCTACAAATGGGTCTCTGGCAATGGAGAGAGCGCGCCGCAAGCCGCCGTC 307
 QY 44 AlaLeuTyLysRgProLy5a1aenG1yValLy5eRserThyValh1sVal1ieser 63

DB 308 GCACTAGCGCGCGGCAATGCAACGAGAGTGAACCCAGATCTACGACCAATCTCC 367
 QY TheProG1a1aSer1y5a1a1ieserCy5yG1yG1n1ieser1yETyRthLeu 83
 DB 64 TTTCTGGGCTACAAATGGGTCTCTGGCAATGGAGAGAGCGCGCCGCAAGCCGCCGTC 307
 QY 368 AACCACTGGTCTCCAAAGCCCGTGAATACCGAGGCGACACAGATCTCATTCACACTG 427
 DB 84 SerTyRaenG1nThyValValValG1uTyRth1sAaRyLysRgThyAaRgPheGln 103
 QY 428 TCCCGAGTCACCTCACTGATGAGTGAACACATGACGACCAAGCATGTTCCAG 487
 DB 104 ValG1yAaRgSerThyG1uSerPro1LeaRgPheValThyAaRgThy1ieserGlyser 123
 QY 488 ATGGCGGCTTACATGAACATGATGATCTTTGGTGAACACATCTCTCTGGTGA 547
 DB 124 G1nAaRThyAaRg1uAaG1n1eThRg1nserThy1ieserAaRgPheA1aG1yPheAaR 143
 QY 548 GGGGCTACA---GAGGCGCTTCTGCGCAAGTACATCTCCGCTATGCTGCAATTT 604
 DB 144 ValCyAaRPaRgAenG1uProTyRThyAaRg1ePheA1aAaG1yPheAaR 163
 QY 605 CTCTGAGACCGCGCGCACCTTAATCTGCCGATCTAGCTGCTGCTGATGCTCT 664
 DB 164 Ly5aen1LePheLeuG1yG1uLy5a1a1y5eRProLy5aenProaRg1y1sMetAaR 183
 QY 665 AGCAATCTTTCTTGAAGAGCGGCGACCAATGAGACCTCCGATGATGATGAC 724
 DB 184 G1yLeuThyThyAaenG1yValLeuValMeTh1sProaRg1yG1yPheThyG1uG1uSer 203
 QY 725 GGGTTACAAACCAATGGGTCTCTGGTGAACGACGAGTGGCTTCCGAGACTCT 784
 DB 204 G1nProG1yVal1TPaRgG1u1ieserValCy5yG1yAaRValTyRThyLeuAaRg1uThr 223
 QY 785 GCGCCAGGTCTGAGAGGAGATTTCTGTCGGGAAATGTGAACGCTGCGGAGTAC 844
 DB 224 ArgSerA1aG1nG1aRg1yLy5aenValG1uSerG1uThyAaRValLeuG1nAaRg1y 243
 QY 845 CAGTACGCTCAGAGGAGGAGGAGGCTGTGTAAGAAAGAAATCCAACTCTGCAAGACG 904
 DB 244 SerLeu1LeaRyLeuCy5yG1yAaThyLeuLeuTPaRgThyA1aAaRg1yLeuPhe1s 263
 QY 905 TGGCTCATGACCTGTGTGGGCGACACTGTGTGGCGGACCTCGGCTGGTTCGTAAG 964
 DB 264 ThyProThyG1nLy5e1LeG1uA1eAaRgG1nG1u1LeaA1aAaRgProG1n 283
 QY 965 GCAACCCACCTAAACCACTAAGGCGCCAGCAGAGAGGCAACCAACAGCGCCGAG 1024
 DB 284 Cy5ProValG1yLeuAaRThyLeuA1aRgProSer1LeaAaRgLy5eG1uVal1G1u 303
 QY 1025 TCCCGGAGGCTTGAACACTTGGCTTCCGACGCCCGGCTGGCGGACAGACCC 1084
 DB 304 G1uLy5---G1nProTPaR1aTyLeuSerCy5G1yAaValh1sG1yTyRThyAaR 322
 QY 1085 GACAAACAGCAACCTGGGCTGACGTCGCTGTGGTCAATGTCATGACGCGCTG 1144
 DB 323 G1yH1sAaRgSerAaRThyG1uA1aAaenG1uAaRgG1uCy5ProMeCyAaRThyValG1y 342
 QY 1145 GCGTCGCGAGGAGAACAGAGCCCGCAGAGGAGTGTCTCTGCGCGCTTGGGGA 1204
 DB 343 ProTyRValProLeuThyLeuG1yCy5yG1yAaRgThyValAaRPaRg1yProPro 362
 QY 1205 CCTATGTCCTCTGTCGCTGAGAGGCGGCTCTGCTGCTGACCTGGGACACCC 1264
 DB 363 Thh1sAaRPhThyProCy5yG1yH1sValCy5SerG1yLy5eR1aLeuTyRThySer 382
 QY 1265 AGCAGAGCTTTTGAACCTGTGGCCACGTCGTTGAGAGACACGCGCGCTAGTGGCT 1324
 DB 383 G1n1LeProLeuProh1sG1yThh1sAaRPh1sAaR1aAa1aCy5ProPheCyAaR1aThr 402
 QY 1325 CAGACACCTCTGCGAGCAGCAGCATGCTTTCCACGCTGCTGCGCTTCTGTGAGCT 1384
 QY 403 G1nLeuValG1yG1nAaenCy5y1eLy5e1LePheG1nG1yPro1LeaR 420

DB	1385	TGGCTCACCGGAGCTTGCTGTCCGCGCTCATTTTCCAGGGCCACTGGAC	1438
RESULT 8	AY408790	1242 bp	DNA
LOCUS	AY408790		linear
DEFINITION	Pan troglodytes HCM341 gene, VIRTUAL TRANSCRIPT, partial sequence.		SSS 12-DEC-2003
ACCESSION	AY408790		
VERSION	AY408790.1	GI:39764758	
KEYWORDS	SSS.		
SOURCE	Pan troglodytes (chimpanzee)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.		
REFERENCE	1 (bases 1 to 1242)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Interferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2 (bases 1 to 1242)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	location/Qualifiers		
source	1..1242		
gene	/organism="Pan troglodytes"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9598"		
	<1..>1242		
	/locus_tag="HCM341"		
ORIGIN			
Alignment Scores:			
Pred. No.:	3.56e-153	Length:	1242
Score:	1626.50	Matches:	298
Percent Similarity:	80.91%	Conservative:	41
Best Local Similarity:	71.12%	Mismatches:	73
Query Match:	71.03%	Indels:	7
DB:	9	Gaps:	4
US-10-041-030-4 (1-420) x AY408790 (1-1242)			
OY	3	SePProgiyngIngluGhHieCyahLaProhAnlysgUProVallyeTyrgIyGluue	22
Db	1	TCTCCCGAT-----GAAGATGCCGACGACGAGCGAGGCCATTCAGTATGGTGAATC	54
OY	23	ValValleuGyTyTAsngIyAlaleuProhAnngIyAaPaRgIyArGArLySeSerAr	42
Db	55	ATCGTCCTGGGCTCAATGGGCTGTGGCAAGTGGCGAAGAGGCCGCCGCGAAGCCG	114
OY	43	PheAlaleuTyTyLbArPProLyAlAaAnGlyVallybProSeThrValHieValile	62
Db	115	CTGGCACTGACGACCGCGCGCCACGCCCAACGGGGTGAAGCCAGACGTCATCACACATC	174
OY	63	SeThrProGInAlaSeTyAlAlleSeCyAlblysgIyngInHieSerTleSeTyThr	82
Db	175	TTCACGCGG-----CTGCGACATGAGTAACCGTGTCAAGACAGCATTCGTATACA	225
OY	83	LeuSeArGaAnGInThrValValValGlyTyThrHieAaPlybAaPThrAspMetPhe	102
Db	226	CTGTCTGGAGACCACTCGTCACTAGTGTGAAGTATACACTGATATACGACAGCATGTTC	285
OY	103	GlnValGlyArGSeThrGlnSeTyroIleAspPheValValThrAspThrIleSeTyLy	122

Db	286	CAGNNNGCCGCTCCACGAGAAACATGATTGACTTCGCGTGAACAGACCGCCNNNTGCA	345
OY	123	SerGlnAenThrAspArgLysAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArg	142
Db	346	GGA---GGGGCTGCCAGAGGGCCCTTCTGCCCGAGACCACTCTCGCTATGCCCTGGCC	402
OY	143	IleValCysAspArgAsnGlnProTyrThrAlaArgIlePheAlaAlaGlyPheAspSer	162
Db	403	ATCCTCTGTGACCGCGCGGCCACCCCTATCTCGCCGCATCTATGCGCGCTTTCGATGCC	462
OY	163	SerIysAsnIlePheLeuGlyGlnLysAlaIleValSerPheLysAsnProAspGlyIleSer	182
Db	463	TCTAGCAACATCTTCTTGGAGAGGAGCGGCCAATATGGCGGACCCCAAGATGGCTTGATG	522
OY	183	AspGlyLeuThrThrAsnGlyValIleLeuValMetHisProArgGlyGlyPheThrGlnGlu	202
Db	523	GATGAGCTGACCACTAATAGAGTCTCTGTATGACCCCGCAGCGGGCTTCTCCAGAGAC	582
OY	203	SerGlnProGlyValTTPArgGlnIleSerValCysGlyAspValTyrThrLeuArgGlu	222
Db	583	TCAAGCCCGGGTGTCTGGCGGGAAGTCTCGTCTGTGGGAATGTGTACACTTGGCGGAC	642
OY	223	ThrArgSerAlaGlnGlnAlaArgGlyLysLeuValGlnSerGlnThrAsnValLeuGlnIle	242
Db	643	AGCGGCTCAGCCACAGACCGGGGGCAAGCTGTGTGAAGAAACGATGCCAAGTCTGCAGAAC	702
OY	243	GlySerLeuIleAspArgLeuCysGlyValAlaThrIleLeuLeuTTPArgThrAlaAspGlyLeuPhe	262
Db	703	GGCTCTCTCACTGACCTGTGTGGGGCCACACTGCTGTGGCCACACCGCGGGCTGCTG	762
OY	263	HisThrProThrGlnLysHisIleGlnAlaLeuArgGlnGlnIleAsnAlaAlaArgPro	282
Db	763	CGGGCTCCCACTGAAGACACTGAGGCGCCAGCGGCGAGAGGCAAAATGCAAGCGGGGCC	822
OY	283	GlnCysProValAlaGlyLeuAsnThrLeuAlaPheProSerIleAsnAlaGlyGlnValAla	302
Db	823	CAGTGCCTCCGTGGGCTTACAGCACTGTGGCTTCCCAAGCCAGCCCGTGGCCGACAGCG	882
OY	303	GlnGluLys---GlnProTTPAlaAlaTyrLeuSerCysGlyHisValHisGlyTyrHisAsn	321
Db	883	CCCGACAAACAGCAACCTGGGTGTACGTCCGCTCGGGGACAGTCCACGGCTTACACAGCG	942
OY	322	TTPAlaHisArgSerAspThrGlnAlaAsnGlnArgGlnCysProMetCysArgThrVal	341
Db	943	TGGGGCTGCGCGCGGAGCGGGGCCCCACAGAGCGCAATGTCCTCTGCGCGCTTGTG	1002
OY	342	GlyProTyrValPProLeuThrPLeuGlyCysGlnAlaGlyPheTyrValAspAlaGlyPro	361
Db	1003	GGGCTTATGTGCTCTATGTGCTTGGCCAGAGGCGCGGCTCTGTGCTGACCTTGGGGCG	1062
OY	362	ProThrHisAlaPheThrProCysGlyHisValCysSerGlnLysSerAlaLysTyrThr	381
Db	1063	CTTAGCCATGCTTTCGACCTTGGGGGCGCACGTCTGCTTGAGAAACATGCGCGCTACTGG	1122
OY	382	SerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCysAla	401
Db	1123	GCCCAAGACACACATGCCACAGCGGACCAACCATCTTTCATGCGGCTGCTTTGGCGGG	1182
OY	402	ThrGlnLeuValGlyGlnGlnAsnCysAlleLysValLeuIlePheGlnGlyProIleAsp	420
Db	1183	GCTGTGCTTACTGCGAGCATGTGCTGGTCCGCTCATTTTCACAGGGCCCGCTGGAT	1239
RESULT 9			
LOCUS	CR606152	2572 bp	mRNA linear HTC 21-JUL-2004
DEFINITION	full-length cDNA clone CSODN004Yg19 of Adult Brain of Homo sapiens (human).		
ACCESSION	CR606152		
VERSION	CR606152.1	GI:50486959	
KEYWORDS	HTC; CNSLT_cDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		

JOURNAL REFERENCE AUTHORS	gene trios Science 302 (5552), 1960-1963 (2003) 14673302 2 (bases 1 to 1242)
	Clairx,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejeriwal,A., Todd,M.A., Tanenbaum,D.M., Givello,D.R., Lu,F., Murphy,B., Perreira,S., Wang,G., Zheng,X.H., White,T.J., Snihsy,J.J., Adams,M.D. and Cargill,M. Direct Submission
TITLE JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES source	location/Qualifiers 1..1242 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>1242 /locus_tag="HCM3341"
ORIGIN	
Alignment Scores:	
Pred. No.:	1.14e-152
Score:	1621.50
Percent Similarity:	80.91%
Best Local Similarity:	70.88%
Query Match:	70.81%
DB:	9 Gaps: 4
US-10-041-030-4 (1-420) x AY408789 (1-1242)	
OY	3 SerProGlyGlnGlnGlnGlnHsiCySaLaProAenLyGlnProValLySerGlyGlnLeu 22
Db	1 TCTCCCGCGT-----GAAGATGGCGACGACGAGGAGAGACCCATGATAGTGAATC 54
OY	23 ValValLeuGlnGlyTyrGlnGlyValAlaLeuProAenGlyYAspArgGlyYArgArgSerArg 42
Db	55 ATCTCTCTGGCGCTTCAACATGGTGTCTGCGCAAGTGGGACAAAGGCGCGCGGGAAGCGC 114
OY	43 PheAlaLeuTyrYAspArgProLySAIAlaGlnGlyValLyPProSerThrValHisValIle 62
Db	115 CTGGCACTGAGACGCGCGGTGACACGCGCAAGGGGTGAAGCCAGATGATGACACACATC 174
OY	63 SerThrProGlnAlaSerLySAIAlaIleSerCySAIyGlyGlnHisSerIleSerTyrThr 82
Db	175 TCCACGCGCG-----CTCGACACGATGAACCGGTGTACGACACAGCATCTCGATAC 225
OY	83 LeuSerArgGlnGlnThrValAlaValAlaGlyTyrThrHisAspLyAspThrAspMetPhe 102
Db	226 CTGTCCCGGAGCCACTGGTCAATAGTGGAGTATACATGATAGGAGACACAGACATGTC 285
OY	103 GlnValAlaGlySerThrGlnSerProIleAspPheValValThrAspThrIleSerGly 122
Db	286 CAGATTGGCGCGCTCCACAGAGAACATGATTGACTTCGTGTACAGACACGTCGCCCTGGA 345
OY	123 SerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAlaCyArg 142
Db	346 GGA---GGGCGCTGGCGAGGGCGCTTGTGCCACAGACACCATCTCCGCTATGGCTGGCGC 402
OY	143 IleValCySAIAspArgGlnGlnProLyTyrThrAlaArgIlePheAlaAlaGlyPheAspSer 162
Db	403 ATCTCTGTGACCGCGCGGACCCCTTATACGTGCCCGGATGTAGCGCGCTTGTGAATGCC 463
OY	163 SerLyAsnIlePheLeuGlyGlyLySAIAlaIleYTPLyAsnProAspGlyHisMet 182
Db	463 TCTAGCAACATCTTCTCTGAGAGGCGACGAGGCGCAATGCGGACCCAGATGGCTGTATG 522
OY	183 AspGlyLeuThrThrAsnGlyValIleValMetHisAspArgGlyGlyPheThrGlnGlu 202
Db	523 GATGACCTGACCAACATGATGATCTCTGTGTGATGACACCGGCGGGGCTTCTCCAGGAGAC 582
OY	203 SerGlnProGlyValTTPArgGlnIleSerValCySAIyAspValYrThrLeuArgGlu 222

Df			583	TCAAGCCCCGGGTGCTCGCCGGAGAATCTCGTCTGTGGAAATGTGCATCATTCGGGGAC	642
OY			223	ThrArgSerAlaGlnGlnAtrGdLYLeuValGlusergluthraenValleuGlnasp	242
Df		:::::	643	AGCGGCTCAGCCACGACGCGGGGCAACTGGTGGAACACAGATGCCAACGCTGCGCACAGAC	702
OY			243	GlySerIleuIleAparLeuCySGLYVALThrLeuLeuTrpArgTrpAlaAspGLYLeuPhe	262
Df			703	GAGCTCTCATCAACCCTGTGTGGGCCACACTGTGTGGGGCACACCGCGGGGCTCTGTG	762
OY			263	HieThrProthrGlnLYSHNslleglnAlaLeuArgGlnGlnulleaenAlalaArgPro	282
Df		:::::	763	CGGGCTCCCACTGGAAGCAACTGAGGCCACGCGGACAGGCAATAATGACGCGGGCCC	822
OY			283	GlnCYarProValIGLYLeuAnThrLeuAlaPheProserIleAsnArgLYSGlnValVal	302
Df		:::::	823	CAGTGGCCCCGTGGGCTCAGACACTGTGCTCCCTCCCAAGCCACGCCCTGGCCGACAGCG	882
OY			303	GluGlnLYS---GlnProTrpAlaTYrLeuSerCYsgLYHISVALHNslgLYTHIASn	321
Df		:::::	883	CCGACAAACAGACGCCCTGGGTCTTAGTCGCGGTGGGGACAGTCATGGCTACAGGCG	942
OY			322	TrpGLYHISArgSerAapThrGlnAlaAsnGlnAArgGLUCYSPrometCYeHGYthrVal	341
Df		:::::	943	TGGGGCTGTCGGCGGCGCGGGCCCCCAGAGGCGCAAAATGCTCTCTCTCGCGCTTTGTG	1002
OY			342	GlyProTYrValProLeuTrPLEuGLYCysGlnAlaGlyPheTYrValAspAlaGlyPro	361
Df		:::::	1003	GGGCTTAATGTGCTCTCATGTGCTTGGCCAGAGGCGCGGCTGTGCTGACCCCTGGGCGG	1062
OY			362	ProThrHisAlaPheThrProCYsgLYHISVALCYsSerGlnLYSeralALyTYrTrp	381
Df			1063	CTTAGCCATGCTTTGGACCTTGCGGCGCACGTGCTGCTAGAAAGACTGCCGCGCTACTGG	1122
OY			382	SerGlnIleProLeuProHISglYTHrHisAlaPheHISalAlaIaCYsProPheCYaLa	401
Df		:::::	1123	GCCCAAGACACACTGCCCCACAGGACCCATGCTTTCATGCGGCTCCCTCTTTGGCGG	1182
OY			402	ThrGlnLeuValIGLYGlnGlnAsnCYSyleLeuIlePheGlnGlnIYProIleasp	420
Df		:::::	1183	GCTGTGGCTTACCGCGGACATGGCTGGTCCGCTCATTTTCCAGGGCCGCGTGGAT	1239
RESULT 11					
LOCUS	AY408791	1242 bp	DNA	linear	GenSeq 12-DEC-2003
DEFINITION	Mus musculus HCM3J41 gene, VIRTUAL TRANSCRIPT, partial sequence,				
ACCESSION	AY408791				
VERSION	AY408791.1	GI:39764759			
KEYWORDS	GSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	(bases 1 to 1242)				
	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,				
	Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,				
	Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Snihsny,D.J.,				
	Adams,M.D. and Carcill,M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous				
	gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCES	(bases 1 to 1242)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,				
	Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,				
	Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Snihsny,D.J.,				
	Adams,M.D. and Carcill,M.				
TITLE	Direct Submision				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,				
	Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering				

them based on alignment.

FEATURES
Source Location/Qualifiers
1.1242
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
gene
<1..>1242
/locus_tag="Hm3341"

ORIGIN

Alignment Scores:

Pred. No.:	1.17e-151	Length:	1242
Score:	1611.50	Matches:	295
Percent Similarity:	81.10%	Conservative:	44
Best Local Similarity:	70.57%	Mismatch:	72
Query Match:	70.37%	Indels:	7
DB:	9	Gaps:	4

US-10-041-030-4 (1-420) x AY408791 (1-1242)

```

QY 4 ProGlyGlnGluGluHisCysAlaProAsnLysGluProValIlyrGlyGluLeuVal 23
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    4 CCGTGTGAAGAA-----GCACGTGCAGAGGAGGAGCCCATCAAGATGTGTAACATCATC 57
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 24 ValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGlySerArgPhe 43
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    58 GTTCTGGGCTACATAGGCTGTCTGGCAAGTGAACAAAGGCGCCGCCAAGCGCGCTG 117
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 44 AlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHisValIleSer 63
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    118 GCACAGGCGCGCGCCGACATGCCACGAGTGAAGCCAGATGTCATGCACACATCTCC 177
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 64 ThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSerTyrThrLeu 83
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    178 ACACCA-----CTCGCCCTGATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 228
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 84 SerArgAsnGlnThrValValValGluTyrThrHisAspLysAspThrAspMetPheGln 103
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    229 TCCCGGAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 288
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 104 ValGlyArgSerThrGlySerProLysAspPheValValThrAspThrIleSerGlySer 123
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    289 ATTGGCGGCTCTACTGAAACATGATGATGATGATGATGATGATGATGATGATGATGATG 348
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 124 GlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIle 143
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    349 GGGGCTACA---GAGGGCCCTTCTGCCCAAGTACATCTCCGCTATGCTGAGATTT 405
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 144 ValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPheAspSerSer 163
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    406 CTCTGTGACCGCGCGGACCCCTATACGCCGGAATCTACCTGCTGTGATGCTCTCT 465
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 164 LysAsnIlePheLeuGlyGluLysAlaAlaLysTyrLysAsnProAspGlyHisMetAsp 183
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    466 AGCAACATCTTCTTGGAGAGCGGGGACCAATGAGAGCTCCGATGCTGATGAGAC 525
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 184 GlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThrGluGluSer 203
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    526 GGGCTGACACCAATGAGGTTCTGTGATGATGATGATGATGATGATGATGATGATGATG 585
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 204 GlnProGlyValIleProArgGluIleSerValCysGlyAspValTyrThrLeuAlaGluThr 223
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    586 GCCCCAGAGTCTGAGAGGAGATTTCTGCTGCGGGAATGTACACGCTGCGGAGTAC 645
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 224 ArgSerAlaGlnGlnArgGlyLysLeuValGlyLysSerThrAsnValIleGlnAspGly 243
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    646 CGATCAGCTACGACCGCGGAGAGCTGTGAGAAACAAATCCAACTCTCTGCAAGAGCT 705
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 244 SerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrThrAlaAspGlyLeuPheHis 263
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    706 TCCCTCATGACCGCTGTGAGGAGCAACATGTTGGGAGCACTCCGAGCTGGGTTGCTGAG 765
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 264 ThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaAlaArgProGln 283

```

```

DB 766 GCACCCACCTGAAACACTAGAGGCCACGAGCAGAGGCCAACGACGAGCCGCCAG 825
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    284 CysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValIleGlu 303
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    826 TGCCCGGTGGGCTTACGACCTTGGCTTCCAGGCCAGCGCGTGGCCGACAGCACCC 885
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 304 GlnLys--GlnProThrAlaTyrLeuSerCysGlyHisValHisGlyTyrHisAsnThr 322
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    886 GACAAACAGACACCCGTGCTACGTCCGTGTGTGTCATGTCATGTCATGTCATGTCATG 945
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 323 GlnHisArgSerAspThrGluAlaAsnGlnArgGluCysProMetCysArgThrValGly 342
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    946 GGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1005
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 343 ProTyrValProLeuThrPheGluGlyCysGluAlaGlyPheTyrValAspAlaGlyProPro 362
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    1006 CCTATGTGCGCCCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 363 ThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyrTrpSer 382
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    1066 AGCCAGGCTTTTTCACCTGTGCGCACGCTGTCTGTGAGAAAGACTCCCGCTACTGGGCT 1125
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 383 GlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCysAlaThr 402
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    1126 CAGACACCGCTGCGGACGAGCCACCATGCTTTCACGCTGCTGCTGCTGCTGCTGCTG 1185
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 403 GlnLeuValGlyGluGlnAsnGlyLysLeuLeuIlePheGlnGlyProIleAsp 420
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
    1186 TGCTCATCCGCTGACCTTGGCTGTGTCCGCTCATTTTCCAGGGGCGACCTGAGAC 1239
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 12
AK033815 2474 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus adult male epididymis cDNA, RIKEN full-length enriched
DEFINITION library, clone:9230114E15 product:pellino 2, full insert sequence.
ACCESSION AK033815 GI:26329492
VERSION AK033815.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED 10349636

TITLE
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED 11042159

AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
Komno, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED 11076861

TITLE
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.

```

TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 2474)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source Location/Qualifiers
1. 2474
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:9230114E15"
/db_xref="taxon:10090"
/clone="9230114E15"
/sex="male"
/tissue_type="epididymis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
2. 854
/note="unnamed protein product; pellino 2 (MGI:1891445, GB|NM_033602, evidence: BLASTN, 99%, match=864)
putative"
/codon_start=3
/protein_id="BAC28485.1"
/db_xref="GI:26329493"
/translation="RPRCRIVCDRNEPYARIFAAGDSSKXNFIQEXAKKXNPNGMGLITNGVIVAHFPGSGFTBESQPGVREISVCGDYTTIRBTRSAQQRKLYESSTNVLDGSLIDLCGATLLMRADGLFHPATQGHIEALRQEIINARPCQGVAGNTLAPPSINRKEVEKQPMATLSCGHVGHYSGHGRBDTEANRECECMCRVGVYVPLWAGCEGFIYDAGPPTHTAFTPCGHVCSSEKAKYWSQIPLPHGTHTFAACPCFATQVGFQNCIKLIFQGPVD"
3. 2451
polyA_signal
4. 2474
polyA_site
/note="putative"
/note="putative"
/note="putative"
ORIGIN
Alignment Scores:
Pred. No.: 2,65e-146 Length: 2474
Score: 1563.00 Matches: 279
Percent Similarity: 99.65% Conservative: 3

Best Local Similarity: 98.59% Mismatches: 1
Query Match: 68.25% Indels: 0
DB: 3 Gaps: 0
US-10-041-030-4 (1-420) x AK033815 (1-2474)
QY 138 ACPheA1aCYaRG1IeValCYaSPARgaenGIuProTYrThraLaRG1IePheA1a 157
DB 3 AGCTTCGATGACATGAGATCGTGTGTGACAGGAACGACCATACAGACGACGATATTCGG 62
QY 158 A1aGIyPheAPSeSerIySaen1IePheLeuGIyLulYVa1aLaYsTrIySaen 177
DB 63 GCAGATTCGATTCATTCACAAAATATCTTTCTTGAGAGAAAGCAGCAAAATGAAAAAC 122
QY 178 ProAPG1yH1eMeAPSG1yLeuThrIySaenGIyLulYVa1aMeH1eProARgIy 197
DB 123 CCGATGACACATGATGATGACTCTACCAACGATGCTGTGTATGACCCGCAAGGA 182
QY 198 GIyPheThrGIuGIuSerGIuProGIyVal1TrPARGIu1IeSerValCYaGIyAPVa1 217
DB 183 GGCCTTCACCGAGAAATCCAGCTGAGTCTGAGAGAGATCTCTGTGTGGATGTA 242
QY 218 TYrThrLeuAG1uThrIySaen1aGIuGIuAG1yLulYVa1eGIuSerGIuThr 237
DB 243 TACACCTTCGAGACACACAGATCGGCCAGACAGAGGAAAGCTGTGAAATGAGACC 302
QY 238 AenVal1IeGIuAG1ySer1IeGIu1IeAPSeuCYaGIyVal1aThrLeuThrPARgThr 257
DB 303 AACGTCCTGCAAGACGCGCTCTCATGACCTGTGTGGGCCACCTCTCTGGAGAAC 362
QY 258 A1aAPSG1yLeuPheH1eThrProThrGIuYH1eGIuVal1aLeuAG1uGIu1Ie 277
DB 363 GCAGATGCGCTTTTTCAGGCTCTACTGACAGACATAGAAAGCCCTCGGACGAGATC 422
QY 278 Aaen1a1a1aAGPProGIuCYaProValGIyLeuAenThrLeuA1aPheProSer1IeAen 297
DB 422 AATGACGCGCAGCCCGACGTCCTGTGGCCCTTAACACCTGCTTCCCGACATCAAC 482
QY 298 ArgIySG1uValGIuGIuYSG1uProThrPA1aTYrLeuSerCYaGIyH1eValH1e 317
DB 483 CGAAGAGAGTGTGAAGAGAGAGCGCTGGGACATCTAGGCTCGGACGATGTGAC 542
QY 318 GIYTYrH1eAenTrIG1yH1eArGSeArPThrGIu1aAenGIuARGIuCYaProMet 337
DB 543 GCGTACACACAGCTGGGGCATCGAGCGACAGAAAGCAAGAGGAGTGTCCATG 602
QY 338 CYaH1eThrValGIyProTYrVal1ProLeuThrPLeuGIyCYaGIuA1aGIyPheTYrVal 357
DB 603 TCGAGGACTGTGGGCGCTTACCTCTGTGGCTGTGAGGCGAGATTTATGTG 662
QY 358 AaPA1aGIyProProThrH1eA1aPheThrProCYaGIyH1eValCYaSerGIuYsSer 377
DB 663 GATGGGGAGCCCAACTCAAGCTTACCCCTCGGGACAGCTGTTCAGAGAAAGTCT 722
QY 378 A1aYsTYrTrPSeGIu1IeProLeuProH1eGIyThrH1eA1aPheH1eA1aCYs 397
DB 723 GCCAAGTACTGTTCAGATTCACCTGCCCCAGAGACGACGCTTACAGCCCGCTGT 782
QY 398 ProPheCYa1aThGIuLeuValGIyGIuInAenCYs1IeYLeu1IePheGIuGIy 417
DB 783 CGTTCGTGGCCACAGCTGTGTGTGAACAGAACTGCATTAATTTATTTCCAAAGT 842
QY 418 Pro1IeAP 420
DB 843 CCAAGTGAAC 851
RESULT 13
BX388547
LOCUS BX388547 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION BX388547 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
ACCESSION BX388547
VERSION BX388547.1 GI:30460950

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Eukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 849)
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7349.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?as=CSOXT0082A06_T0713_1&c=7349.f.
Location/Qualifiers
1. 849
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D006YF18"
/issue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
Pred. No.: 1.44e-142 Length: 849
Score: 1519.00 Matches: 272
Percent Similarity: 98.56% Conservative: 1
Best Local Similarity: 98.19% Mismatches: 4
Query Match: 66.33% Indels: 0
DB: 5 Gaps: 0

US-10-041-030-4 (1-420) x BX388547 (1-849)

QY 140 AAlaCyAaRgIleValCyAaPaRgAaNgIuProTyRThrAlaRgIlePheAlaAlaGly 159
1 GCCTGAGATGTCGTGTCGACAGNAATGAACCTTACACACACGAGATTCGCCGCCGA 60
QY 160 PheAaPSeSeRtYsAaNIlePheLeuGlyGluYsAlaAlaYsTrpYsAaPProAaP 179
61 TTGAGACTCTTCAAAAACATATTTCTTGAGAAAAGCAGCAAAAGTGAAAAACCCCGAC 120
QY 180 GlyHMeRSeAaPglyLeuThrTrpAaNgIyValLeuValMeThsPSeAaRgIyGlyPhe 139
121 GGGCAGACATGATGGCTCCTACTACTAATGCGCTGTGTGATGCTCCACAGGGGGCTTC 180
QY 200 ThrGluGluSeRgInProGlyValATrPAGluIleSeRValCySgIyAaPValYTr 219
181 ACCGAGAGATCCACGCCCGGGGCTGCGGAGATCTCTGTGTGAGATGTGTACACC 240
QY 220 LeuAaRgIuThrAaRgSeRAlaGInGlnAaRgIyLysLeuValGluSeRgIuThrAaVal 239
241 TTTCGAGAAACCAAGCTCGCGCCAGCAACGAAAGCTGTGAAAGTGAGACCAACGTC 300
QY 240 LeuGlnAaPglySeRleuNIleAaPLeuCySgIyAlaThrIleuSeRTrpAaGTrpAaAaP 259
301 CTGCAAGACCGGCTCTCTCATTTGACTGTGTGGGGCCCTCTCTCTGAGAACGAGAT 360
QY 260 GlyLeuPheHISThrProThrGlnLysHISAlaGlnAlaLeuAaRgInGlnIleAaAla 279
361 GGGCTTTTTCATATCTCAACTCAGAAACATAGAAACCCCTCGCGGAGGAGATTAAAGCC 420

QY 280 AAlaRgProGlnCyAaProValGlyLeuAaThrLeuAlaPheProSeRtleAaRgIy 299
421 GCCCGGCTCAGATGTCCTGTGGGGCTCAACACCTTGCCCTCCAGCATCAACAGAAA 480
QY 300 GluValValGluGluYsGInPSeRTrpAlaYTrLeuSeCySgIyHISValHISGlyTr 319
481 GAGGTGTGAGAGAAAGACAGCCCTGGCATATCTCATGTTGTGGCCAGTCGACGGGATC 540
QY 320 HisAaNTpGlyHISaRgSeRSeRThrGlnAlaAaNgIuAaRgIuCyAaPSeRVal 339
541 CACAACTGGGGCCATCGAGATGACACGAGGCCAAGAGAGGAGATGCCATGTGCAGG 600
QY 340 ThrValGlyProTyRAlaProLeuTrpGluGlyCySgIyAlaGlyPheTyRValAaPAla 359
601 ACTGTGGGCCCCATATGCTCTCTGCTGAGAGAGAGATTTATGTAGACGA 660
QY 360 GlyProProThrHISAlaPheThrProCySgIyHISValCySgIySeRAlaLys 379
661 GAGCCGCCAATCATATCTTCACTCCCTGTGACACGTGTGCGAGAAAGTGTGAAA 720
QY 380 TyTrpSeRgInIleProLeuProHISGlyThrHISAlaPheHISAlaAlaCyAaPProPhe 399
721 TACTGATCTCAGATCCCGATTCCTCATGGAACATCATGATTTACCGCTGCTCCCTTC 780
QY 400 CyAAlaThrGlnLeuValGlyGluGlnAaCySgIleLysLeuAlaPheGln 416
781 TGTGTACACAGCTGAGTGGGAGCAAACTGCATCAATATTTTTCAAA 831

RESULT 14
AK029586
LOCUS
DEFINITION
MUS musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:4930404N17 product:ipellino 1, full insert sequence.
ACCESSION
VERSION
AK029586.1 GI:26081414
KEYWORDS
HTC; CAP trapper.
SOURCE
MUS musculus (house mouse)
ORGANISM
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
METH. ENZYMOL. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636

REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
GENOME RES. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159

REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, R., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsuki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
GENOME RES. 10 (11), 1757-1771 (2000)
MEDLINE
20530913
PUBMED
11076861

REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
NATURE 409, 685-690 (2001)
REFERENCE
5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3531)

AUTHORS Aduchi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku, A., Tanaka, Y., Tanaka, T., Tomari, A., Toya, T., Yasunishi, A., Yamanaka, M., and Hayashizaki, Y.

TITLE Direct Submision

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

FEATURES

source

1.3531

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM:DB:4930404N17"

/db_xref="taxon:10090"

/clone="4930404N17"

/sex="male"

/issue_type="testis"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

1.3531

/note="pe11ino 1 (MGD|MG1:1914495, GB|NM_023324, evidence: BLASTN, 99%, match=2228)"

ORIGIN

misc_feature

misc_feature

Alignment Scores:

Pred. No.: 5,64e-137 Length: 3531

Score: 1473.00 Matches: 271

Percent Similarity: 85.75% Conservative: 42

Best Local Similarity: 74.25% Mismatches: 50

Query Match: 64.32% Indels: 3

DB: 3 Gaps: 0

US-10-041-030-4 (1-420) x AK029586 (1-3531)

QY 58 ThrValHleValIleSerThrProGlnAlaSerIleAlaIleSerCysIysGlyGlnHis 77

Db 2350 ACTGCAATATTTGCTGCTTAACCTTTCATTCATAGCAATAGCAACAGACCGACAT 2409

QY 78 SerIleSerTyThrIleuSerAArganGlnThrValValGluTyThrHisAepLys 97

Db 2410 ACCATATCATATATCTTATCTCAGCCAGACGCGTGGTGAATATATCTCATGACAGC 2469

QY 98 AspThrAspMetPheGlnValGlyArgSerThrGlnSer-ProIleAspPheValAla 117

Db 2470 AACCACTGATATGTTTCAATGCTGCTCAACTGAAAGTCGCTATTGATTTGTATGTAAG 2529

QY 117 rAspThrIleSerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSe 137

Db 2530 TCACACCGCTGCTGGAAAGTCAGAGTAATTCGACACGAGTACAGTAACAAGACATATATC 2589

QY 137 rArgPheAlaCysArgIleValCysAspArgAnGlnProTyThrAlaArgIlePheAl 157

Db 2590 AAGATTGGCTGTAGATCATATGTGACGCGAGTCCCTTTACAGCTCGGAGTTATATC 2649

QY 157 aAlaGlyPheAspSerSerIleValPheLeuGlyGluValAlaAlaTyThrLysAs 177

Db 2650 TCCAGGGGTGTATATCAAAAACATCTTCTTGAGGAAGGCTGCGCAAGTGAAGAC 2709

QY 177 nProAspGlyHisMetAspGlyLeuThr-Thr-AsnGlyValLeuValMetHisProArg 197

Db 2710 ATGTATGGCGCAGATGATGCTTACACAGCAATAGAGTTCTTGTGATGATCCAGTA 2769

QY 197 lGlyPheThrGlnGlnSerGlnProGlyValTyPArgGlnIleSerValCysGlyAsp 217

Db 2770 ATGGGTTCCAGAAACATCCAAACGTGGAATGTGAGAAATATCAGTATGTGGAAACG 2829

QY 217 aITyThrIleuArgGlnThrArgSerAlaGlnGlnArgGlyValLeuValGlnSerGln 237

Db 2830 TCTTAGGTGTGTAAGCCAGATCAGTCAGCAGAGAAAGATGTGAAATTTGAA 2889

QY 237 hrAsnValLeuGlnAspGlySerLeuIleAspLeuCysGlyValaThrIleuLeuTyPArg 257

Db 2890 CCAATCAGGTACAAATGCTCTTAAAGTACCTTTGTGTGCAACCGCTGTGGCGTA 2949

QY 257 hrAlaAspGlyLeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGln 277

Db 2950 CTGCAGAAAGGCTTCCCACTACTGCTAGTCAACACTTAAACCTTAAACAGCGAGA 3009

QY 277 lAsnAlaAlaArgProGlnCysProValGlyLeuAsnThrIleuLapheProSerIleA 297

Db 3010 TTAAGCGGCTGCGCCAGTGCCTGTAGGTTCAACACAGTACCTTCCACAGATGA 3069

QY 297 snArgLysGlnValValGlnGlnLysGlnProTyPAlaTyThrIleuSerCysGlyHisValH 317

Db 3070 AAGAGAGAGATGTTTGAATTAAGTAAGCAACATGGGTATATTAACGCGGCATGTCC 3129

QY 317 lGlyTyThrIleAsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProM 337

Db 3130 ATGCTTATCATTAACGCGGAAACAAAGAAACGACGCGCAAAAGATCGTGAATGCCCTA 3189

QY 337 eCysArgThrValGlyProTyPValProLeuTrpLeuGlyCysGlnAlaGlyPheTyTy 357

Db 3190 TGTGTAGGTGTGTGTCTTATGTCCTCTGTGGCTTGAATGTGAAGCTGAATTTTATG 3249

QY 357 aAlaAspAlaGlyProProThrHisAlaPheThrProCysGlyHisValCysSerGlnLys 377

Db 3250 TGGAGCGCGGCGCTCCACCCATGCTTTAGCCCTGTGGGCACGTGTTCAGAAAGAC 3309

QY 377 eAlaLysTyTrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIAC 397

Db 3310 A-ACGGCTTACGTCGCCAGATCCCATCTCCATGATGTAACGACACTTTCATGACGCT 3368

QY 397 yEProPheCysAlaThrGlnLeuValGlyGlnGlnAsnCysIleLysLeuIlePheGln 417

Db 3369 GCCCTTCTGTCAATCATGTTGGCTGTGAACAAAGGCTATATACAGACTTATTTCCAG 3428

QY 417 lYProIleAsp 420

Db 3429 GACCTTTAGAC 3439

RESULT 15

CN329888 957 bp mRNA linear EST 14-APR-2004

LOCUS CN329888/1

DEFINITION AGECORRT 22003665 NICHD XCC Te2 Xenopus laevis cDNA clone

IMAGE:7207690 3', mRNA sequence.

ACCESSION CN329888.1 GI:46387524

VERSION CN329888.1

KEYWORDS EST.

SOURCE Xenopus laevis (African clawed frog)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2004, 13:12:47, Search time 7535 Seconds
(without alignments)
2635.922 Million cell updates/sec

Title: US-10-041-030-4
2230
Perfect score: 1 MFSPQGEHCAKPKPKVYK.....ATQVGEONCKLIRFGPID 420
Sequence:

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n model -DEV=x1p
-CGEN2.1/USPFC spool.p/US10041030/runat.09122004.095517.7663/app.query.fasta_1.583
-DB=GenEmbl -QFRT=fastcdp -SUPFIX=rgc -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -STRAT=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-OCCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10041030 CGEN 1.1 7406 @runat.09122004.095517.7663 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ha:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_ey:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2230	100.0	1565	6	AX817203 Sequence
2	2230	100.0	1811	9	BC009476 Homo sapi
3	2230	100.0	5579	6	BD190328 Etk1 phos
4	2230	100.0	5597	9	AF302502 Homo sapi

5	2230	100.0	5921	6	AX68405	AX68405 Sequence
6	2230	99.7	1263	6	AR48358	AR48358 Sequence
7	2230	99.7	1263	6	AX298079	AX298079 Sequence
8	2175.5	95.0	1260	6	AR483567	AR483567 Sequence
9	2175.5	95.0	1260	6	AX298077	AX298077 Sequence
10	2138.5	93.4	1646	5	BC072891	BC072891 Xenopus 1
11	2129.5	93.0	1936	5	BC063200	BC063200 Xenopus t
12	2111.5	92.2	1717	6	BD190327	BD190327 Etk1 phos
13	2111.5	92.2	1734	10	AF302504	AF302504 Mus muscu
14	2006	87.6	2470	5	BC075973	BC075973 Danio rer
15	1917	83.7	1257	6	AR483566	AR483566 Sequence
16	1917	83.7	1257	6	AX298075	AX298075 Sequence
17	1917	83.7	1257	6	AF300987	AF300987 Homo sapi
18	1917	83.7	1257	6	AX817197	AX817197 Sequence
19	1917	83.7	3502	9	BC063611	BC063611 Homo sapi
20	1917	83.7	3511	9	BC050533	BC050533 Homo sapi
21	1917	83.7	3526	6	AX817195	AX817195 Sequence
22	1917	83.7	3571	9	BC050019	BC050019 Homo sapi
23	1917	83.7	3571	9	AX68403	AX68403 Sequence
24	1917	83.7	7136	6	AF302505	AF302505 Homo sapi
25	1917	83.7	7136	6	AX817189	AX817189 Sequence
26	1917	83.5	1257	9	HS427859	AJ278859 Homo sapi
27	1917	83.5	1257	6	AR483565	AR483565 Sequence
28	1917	83.5	1257	6	AX298073	AX298073 Sequence
29	1912	83.5	1912	10	BC016515	BC016515 Mus muscu
30	1909	83.4	3710	10	AF302503	AF302503 Mus muscu
31	1884	82.3	3200	5	BC044117	BC044117 Xenopus 1
32	1873	81.8	259038	2	AC128566	AC128566 Rattus no
33	1873	81.8	259038	2	AC119714	AC119714 Rattus no
34	1854	81.0	1732	10	BC027062	BC027062 Mus muscu
35	1822	79.6	3846	6	AX817191	AX817191 Sequence
36	1743	76.1	954	6	CQ718825	CQ718825 Sequence
37	1732	75.6	229573	2	AC109960	AC109960 Rattus no
38	1724	74.8	1126	6	CQ720116	CQ720116 Sequence
39	1642	71.7	1338	6	AR483571	AR483571 Sequence
40	1642	71.7	1338	6	AX298083	AX298083 Sequence
41	1642	71.7	1338	9	AF487457	AF487457 Homo sapi
42	1642	71.7	2631	9	BC025723	BC025723 Homo sapi
43	1640	71.6	2589	6	AX817201	AX817201 Sequence
44	1632	71.3	1933	10	BC028931	BC028931 Mus muscu
45	1632	71.3	1936	6	AX817199	AX817199 Sequence

ALIGNMENTS

RESULT 1	AX817203	1565 bp	DNA	linear	PAT 10-DEC-2003
LOCUS	AX817203				
DEFINITION	Sequence 15 from Patent WO0221138.				
ACCESSION	AX817203				
VERSION	AX817203.1	GI:39722594			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	Schneider, A., Hiemisch, H., Rosener, M., Klugmann, M., Naim, J., Eisenhardt, G., Kuner, R., Lanahan, A., Worley, P., Spielvogel, D. and Schek, S.				
TITLE	The P30 gene family and the utilization thereof				
JOURNAL	Patent WO 02/21138-A1 15-14-MAR-2002;				
FEATURES	Location/Qualifiers				
source	1..1565				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
	51..1313				
	/note="unnamed protein product; codexender Bereich der M33 cDNA des Menschen"				
	/codon_start=1				
	/protein_id="CAB84330.1"				

/db xref="GI:39722595"
 /translation="MPSRQEFHCAPNKEPYKGYELVULGYNGALPNDRGRKSRFA
 LYKRPKANKVRSSTVHVISTPQASAKISCKGHSITYLSNRQIVNEVETHKDTDMF
 QVGRSTESPIDFVVDITISGQNTDEAQITOSTISRFACRIVCDNEPYTAIIFAAG
 DSSKNIPLGEEKAKWKNPDGNDLITTINGHLYMHPRGFTBESQGVWEISVCDVY
 TLRETRSAQORGLKVESETNVLODGLIDLCATLMLRTADLFTPTQKIEALROE
 INAAPOCVGLANTLAFPSINKEVVEESQPAVYSCGHVHNHMGHSDPEANERE
 CPMCRVTGYPVPLMTJGCEAGFPVDAGPPTHAFTPGAHNCSKSAKYMSQIPLPHGTHA
 FPAACPFCAITQVBEONCKILFQGRPID"

ORIGIN

Alignment Scores:

Pred. No.:	3, 87e-177	Length:	1565
Score:	2290.00	Matches:	420
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-041-030-4 (1-420) x AX617203 (1-1565)

QY 1 MetPheSerProGlyGlnGluGluHisCyValaProAenlyeGluProVallyeTygIy 20
 DB 51 AGTGTTCCTCCCTGCGCAGAGAGAACACTGCGCCCAATPAAGAGCAGTGAATACGGG 110
 QY 21 GtuleuValIvalIeuGlyTyraenGlyAlaIeuProAenGlyAspAarglyIyrgIyIs 40
 DB 111 GAGCTGTGTGTCTGTGGGTACATAGTGTCTTACCCATAGAGATAGAGACGAGAGAAA 170
 QY 41 SerArgPheAlaIeuTyrlYsArgProLyAlaenGlyVallyeProSerThraValHis 60
 DB 171 AGTACATTTGCCCTCTCAAGAGGGGCCCAAGGCAATGGTGTCAACCCAGCACCGTCCAT 230
 QY 61 ValIleSerThProGlnAlaSerlyAlaIleSerCySlyeGlyGlnHisSerIleSer 80
 DB 231 GTGATATCCACCCGCCAGCATCAAGGCTATCAGCTGCAAGAGTCAACACAGTATATCC 290
 QY 81 TyrlThreSerArgAsnGlnThraValIvalIeuTyrlThraHisAspLyAspThraP 100
 DB 291 TACACTTGTCTCAAGGATACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 350
 QY 101 MetPheGlnValIyArgSerThrgIuSerProIleAspPheValIyThraSpThrIle 120
 DB 351 AGTGTTCAGGTGGGAGATCAACAGAAACCTTATGATCTTGTGTGTGTGTGTGTGTGTGTGT 410
 QY 121 SerGlySerGlnAsnThraSpGluAlaGlnIleThrgInSerThrIleSerArgPheAla 140
 DB 411 TCTGCAGCCAGAACACGAGCAGAGCCAGATCACACAGACACCATATCCAGGTTCGCC 470
 QY 141 CySArgIleValIyAspAargAsnGluProTyrlThraAlaArgIlePheAlaIagIyPre 160
 DB 471 TGCAGATGT 530
 QY 161 AspSerSerLyAsnIlePheIeuGlyGluLyAlaAlaIySerTrLyAsnProAspGly 180
 DB 531 GACTCTTCCAAAACATATTCTTGTGAGAAAAGGAGCAAGTGTGAAAAACCCGACGGC 590
 QY 181 HisMetAspGlyLeuThraenGlyValIleuValMetHisProArgIyGlyIyPheThr 200
 DB 591 CACATGAGTGGGCTCACTATATGCGCTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 650
 QY 201 GtIuGluSerGlnProGlyValIyTrArgGluIleSerValIyGlyAspValIyThraIeu 220
 DB 651 GAGGAGTCCACCCCGGGGT 710
 QY 221 ArgGluThraSerAlaGlnIyAlaGlyLySleuValIyGluSerGlnThraenValIeu 240
 DB 711 CGAGAAACAGGTCCGGCCACGACAGAGAAAAGCTGTGTGAAAAGTGTAGCAACGTCTCG 770
 QY 241 GlnAspGlySerLeuIleAspLeuCySgIyAlaThraIleuThraIyThraIyAspGly 260
 DB 771 CAGGACGGGTCCCTCATTTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 830

QY 261 LeuPheIsthrProThrGlnIyHisIleGluAlaIeuArgGlnGluIleAsnAlaIa 280
 DB 831 CTTTTCATACCTCCACACAGACATPAGAAGCCCTCCGACAGATTTAACCCGCC 890
 QY 281 ArgProGlnCySProValIyIyLeuAsnThraIleuAlaIeProSerIleAsnArglyGlu 300
 DB 891 CGGCTCATGTGCTGTGGGGGTCAACACCCGTGCTTCCGACATCAACAGAAAGG 950
 QY 301 ValIyAlIyGluIyGlnProTrPAlaTyrlIeuSerCySgIyHisIvalHisIyThHis 320
 DB 951 GTGGTGTGAGAGAAACACCCCTGGGCAATATCTCACTTGTGTGTGTGTGTGTGTGTGTGTGT 1010
 QY 321 AsnTrpGlyHisArgSerAspThrgIuAlaenGluArgGluCySProMetCySArgThr 340
 DB 1011 AACTGGGGCCATCGAGTGAACAGAGGCCCAAGAGAGAGTGTCCATGTGACGACT 1070
 QY 341 ValGlyProTyIyValIyProIeuTrPleuGlyCySgIyAlaIyPheTyValIyAspAlaGly 360
 DB 1071 GTGGGCCCTTATGTCTCTGT 1130
 QY 361 ProProThraIleAlaPheThProCySgIyHisIvalIySerGlnIySerAlaIyTy 380
 DB 1131 CGGCCACTCATGTCTTCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1190
 QY 381 TrpSerGlnIleProIeuProHiaGlyThraHisAlaPheHisAlaIyCySProPheCyS 400
 DB 1191 TGTCTTCAGATCCCTTGTCTTCACTGAACTCATGATTTACGCTGTGCTTGTCTTGTCTTGT 1250
 QY 401 AlaThrgInIeuValIyGlnGluIleAsnCySlyeIyLeuIlePheGlnIyProIleAsp 420
 DB 1251 GCTACACAGCTGT 1310

RESULT 2
 BC009476 1811 bp mRNA linear PRI 29-JUN-2004
 LOCUS Homo sapiens pellino homolog 2 (Prospol1a), mRNA (cDNA clone
 DEFINITION MG:15066 IMAGE:3942712), complete cds.
 ACCESSION BC009476
 VERSION BC009476.2 GI:33872093
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1811)
 Straube, R., D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Klausner, R.D., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marnett, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stadelman, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Ueda, T.B., Toshikuni, S.,
 Abramson, R.D., Mullaly, S.J., Bosak, S.A., Loquillano, N.A., Peters, G.J.,
 McKernan, K.J., Malek, J.A., Gunatane, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hilyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Sanchez, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Fanhey, J., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywicki, M.I., Skalka, U., Smilins, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.,
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 TITLE JOURNAL
 PUBMED 12477932
 AUTHORS Straube, R.
 DIRECT SUBMISSION
 TITLE JOURNAL
 JOURNAL Submitted (25-JUN-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590.

DEFINITION Elkl1 phosphorylation related gene.
ACCESSION BD190328
VERSION BD190328.1 GI:33000067
KEYWORDS MO 03008589-A/68.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Matsuzaki, O., Matsuda, A., Nagano, Y. and Suzuki, N.
Elkl1 phosphorylation related gene
Patent: WO 03008589-A 68 30-JAN-2003;
JOURNAL ASAHI KASEI CORP, OSAMU MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI

COMMENT OS Homo sapiens (human)
PN MO 03008589-A/68
PD 30-JAN-2003
PE 15-JUL-2002 MO 2002JP007174
PR 18-JUL-2001 JP 01P 218204,31-AUG-2001 JP 01P 263450 PR
21-JAN-2002 JP 02P 012176
PI OSAMU MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI PC
C12N15/54, C12N9/12, G01N33/50, G01N33/15, C07K16/40, A61K39/385, PC
A61K31/711

FEATURES
source location/Qualifiers
1..5579
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.8e-176 Length: 5579
Score: 2290.00 Matches: 420
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-041-030-4 (1-420) x BD190328 (1-5579)

QY 1 MetPheSerProGlyGlnGluGlnHisCyValAProAsnLysGluProValLysTyrGly 20
DB 177 AIGTTTCCCTGGCCAGAGGAACCTGCGCCCAATTAAGGAGCCAGTGAATACGGG 236
QY 21 GJLueuValJalIueuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgLys 40
DB 237 GAGCTGCTGCTGCTCGGATCAATGCTTACCCCAATGAGATAGAGAGAGAGAGAA 296
QY 41 SerArgPheAlaLeuTyrTyrAspArgProLysAlaAsnGlyValLysProSerThrValHis 60
DB 297 AGTAGATTGGCCCTTCAACAGCGGCCCAAGGCAAAATGCTGCAACCCAGCACCGTCCAT 356
QY 61 ValIleSerThrProGlnAlaSerLysValIleSerCysValSerGlyGlnHisSerIleSer 80
DB 357 GTGATATCCAGCCGCCAGGATCCAGGCTATCAGCTGCAAGAGTCAACACAGTATATCC 416
QY 81 TyrThrLeuSerArgAsnGlnThrValJalIueuGlyTyrThrHisAspLysAspThrAsp 100
DB 417 TACACTTGTCAAGAAATCAAGACTGTGTGTGTGAGTATACACATCATTAAGATTCGAT 476
QY 101 MetPheGlnValGlyArgSerThrGluSerProLysAspPheValValThrAspThrIle 120
DB 477 AGGTTTCAGGTGGCGGATCAACAGAAAGCCCTATGACTCTGTGTGTCAACAGACGATT 536
QY 121 SerGlySerGlnAsnThrAspGluJalIueuGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 537 TTGCGACGCAACACCGAGCCAGATCACAAGACCAACATATCCAGGTTCCGCC 596
QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160

DB 597 TGCAGATCTGTGGACAGAAATGAACTTACACAGACGGATATTGCGCGCGGATTT 656
QY 161 AppSerSerLysAsnIlePheLeuGlyGlyValAlaIleValTyrPheAsnProAspGly 180
DB 657 GACTCTTCCAAAACATATTTCTTGAGAAAAGGACAGCAAGTGTGAAAAACCCCAACGCG 716
QY 181 HsMeLapArgLysLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
DB 717 CACATGATGGGCTCACTATATGAGCTCTGTGTGTGATGATTCACAGAGGGGCTTCAC 776
QY 201 GJLueuSerGlnProGlyValIleTyrArgGluIleSerValCysGlyAspValTyrThrLeu 220
DB 777 GAGGATCCACAGCCCGGGGTGTGGCGGAGATCTGTGTGTGAGAGTGTATACCTTG 836
QY 221 ArgGluThrArgSerAlaGlnGlnIleArgGlyLysLeuValGluSerGluThrAsnValLeu 240
DB 837 CGAGAAACCAAGTCCGCCACAGCAAGAGAAAGCTGTGAAAGTGAAGCCAACTCTCG 896
QY 241 GlnAspGlySerLeuIleAspLeuCysGlyValAlaThrLeuTyrArgThrValAspGly 260
DB 897 CAGACGGCTCTCTATTTGACTGTGTGGGCTCTCTCTGAGAAACAGCAAGTGGG 956
QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnIleAsnAlaIle 280
DB 957 CTTTTCATATCTCCAACTCAAGACACATTAAGACCTCCGCGAGAGATTAAAGCCGCC 1016
QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
DB 1017 CGGCTCATGTCTCTGTGGGCTCAACACCTGCGCTTCCAGCATCAACAGGAAAG 1076
QY 301 ValValGlnGluLysGlnProTyrPalaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
DB 1077 GTGTGAGAGAACACAGCCCTGGGCAATATCTCACTTGTGGCACGTCAAGCGGATCAC 1136
QY 321 AsnTyrGlyHisArgSerAspThrGlnAlaAsnGluArgGluCysProMetCysArgThr 340
DB 1137 AACTGGGGCCATCCGAGTGAACCGGACCAACGAGAGGATGCCATGTGAGACT 1196
QY 341 ValGlyProTyrValProLeuThrLeuGlyCysGlnAlaGlyPheTyrValAspAlaGly 360
DB 1197 GTGGGCCCCCTATGTCCCTCTGCTGCTGTGGGAGGAGATTTTATGTAGAACCAAGA 1256
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
DB 1257 CGGCACATCTCACTTCTCACTCTCTGTGACACGTTCTGTGGAGATCTGCAGAAATAC 1316
QY 381 TyrSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
DB 1317 TGGTCTCAGATCCGTTGCCCTCATGAACTCATGATTTCACGCTTGCCCTTCTGT 1376
QY 401 AlaThrGlnLeuValGlyGlnGlnAsnCyAlleLysLeuIlePheGlnGlyProIleAsp 420
DB 1377 GCTACACAGCTGTGTGGGAGCAAAATGCAATTAATTTTCCAAAGTCCCAATTAC 1436

RESULT 4
AF302502 5597 bp mRNA linear PRI 05-JUL-2001
DEFINITION Homo sapiens pellino 2 (PELI2) mRNA, complete cds.
ACCESSION AF302502
VERSION AF302502.1 GI:10242352
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Reisch, K., Jockusch, H. and Schmitt-John, T.
Assignment of homologous genes, Pell1/PELI1 and Pell2/PELI2, for
the Pelle adaptor protein Pellino to mouse chromosomes 11 and 14
and human chromosomes 2p13.3 and 14q21, respectively, by physical
and radiation hybrid mapping
JOURNAL Cytogenet. Cell Genet. 92 (1-2), 172-174 (2001)
MEDLINE 21203570

PUBMED 11306823
 REFERENCE 2 (bases 1 to 5597)
 AUTHORS Resch, K., Jockusch, H. and Schmitt-John, T.
 TITLE Direct Submission
 JOURNAL Submitted (05-SEP-2000) Developmental Biology and Molecular Pathology, University of Bielefeld, Universitaetsstrasse 25, Bielefeld, NRW 33615, Germany

FEATURES
 source Location/Qualifiers

1..5597
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 gene .5597
 CDS .1439
 /gene="PEL12"
 /note="similar to Drosophila melanogaster Pellino"
 /codon_start=1
 /product="pellino 2"
 /protein_id="AA015390.1"
 /db_xref="GI:10242353"

ORIGIN
 Alignment Scores:
 Pred. No.: 1..8e-176 Length: 5597
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-041-030-4 (1-420) x AF302502 (1-5597)
 1 MetPheSerProGlyngIngluHh1sCyalaProaenlyegluProvalylyrGly 20
 177 ATGTTTCCCTGGCCAGGAGGAACTGCGCCCAATTAAGAGCCAGGAAATACGGG 236
 21 GluLeuValValLeuGlyTyraenGlyAlaLeuProaenGlyAspArgGlyArGly 40
 237 GAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 296
 41 SerArgPheAlaLeuTyryLybaArgProLybaAlaenGlyVallybProSerThrValHh 60
 297 AGTAGATTGGCTCTCAAGAGCGCCCAAGGCAAAATGATGTCAACCCAGACCGTCAT 356
 61 ValIleSerThrProGlnAlaSerLybaAlaIleSerCybaLybGlyGlnHhSerIleSer 80
 357 GGTATATCCAGCGCCAGGATCAAGGCTATACGCTGCAAGAGTCAACACATATATATCC 416
 81 TyrThleuSerArgenGlnThrValValValGlyTyryThhAspLybaPThrAsp 100
 417 TACACTTGTCAAGAAATCAGACTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 476
 101 MetPheGlnValGlyArgSerThrgLusSerProIleAspPheValValThraPThrIle 120
 477 ATGTTTCAAGTGGGAGATCAACAGAAAGCCCTATCACTTCTGTGCACAGACAGATT 536
 121 SerGlySerGlnSerThraPThrValValValGlyTyryThhAspLybaPThrAsp 140
 537 TCTGGAGCCAGAACCGGAGGAGCCAGATACACAGGACCAATATCAGATTCGCGC 596
 141 CybaArgIleValCybaPArgenGluProTyryThrAlaArgIlePheAlaIleGlyPhe 160
 597 TGCAGATGCTGTGGCCAGGATGAACCTTACACAGCACGATATTCGCGCGCGGATTT 656

161 AspSerSerLybaenIlePheLeuGlyGlyValAlaAlaIleGlyTrpLybaenProAspGly 180
 657 GACTCTTCCAAATAATATTTCTGGAGAAAGGCGCAAGTGGAAAAACCCGACGGC 716
 181 H1sMetArgGlyLeuThrThraenGlyValLeuValMetHh1ProArgGlyGlyPheThr 200
 717 CACATGATGGGCTCACTCTAATAGCGCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 776
 201 GluGluSerGlnProGlyValIleTrpArgGlyIleSerValCybaLybaPThrIle 220
 777 GAGAGATCCACCGCGGGTCTGGCGGAGATCTGTGTGTGGAGATGTGTACACTTGG 836
 221 ArgGluThrArgSerAlaGlnGlnArgGlyLybLeuValGluSerGluThrAsnValLeu 240
 837 CGAGAAACAGATCGCGCCAGCAGACAGAGAAAGCTGTGGAAATGTAGACCAACGCTCG 896
 241 GlnAspGlySerLeuIleAspLeuCybaGlyValIleThrLeuThrArgThrAlaAspGly 260
 897 CAGAGCGCTCCCTCTTAATACCTGTGTGGGCGCACTCTCTTGGAGAACAGCAGATGGG 956
 261 LeuPheHh1SerThrProThrgInlybHh1sGlyAlaLeuArgGlnGluIleAsnAlaIle 280
 957 CTTTTTCATACCTCAACTCAGAAAGCAGATAGAAGCCCTCGGACAGAGATTAAACGGCCG 1016
 281 ArgProGlnCybaProValGlyLybaenThrLeuAlaPheProSerIleAsnArgLybGly 300
 1017 CGGCTCACTGCTCTGTGGGGCTCAACACCTGGCGCTTCCAGCATCAACAGGAAAGG 1076
 301 ValValGluGluGlyLeuProThrgAlaTyryLeuSerCybaGlyHh1sGlyTyryHh1s 320
 1077 GTGGTGGAG 1136
 321 AsnTrpGlyHh1sArgSerAspThrgAlaAlaenGluArgGlyCybaPheMetCybaArgThr 340
 1137 AACTGGGCGCATCGAGTGACACGAGGCGCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1196
 341 ValGlyProTyryValProLeuThrLeuGlyCybaGlyAlaGlyPheTyryValAspAlaGly 360
 1197 GTGGGCGCTATGTGCTCTGTGGCTGTGGCTGTGGAGGAGATTTATGTAGACGACAGA 1256
 361 ProProThrHh1AlaPheThrProCybaGlyHh1sValCybaSerGlybSerAlaLybTyry 380
 1257 CGGCCACATCACTCTTCACTCTCTGTGGACACGATGCTCGAGAGAGCTCGCAAAATAC 1316
 381 TrpSerGlnIleProLeuProHh1sGlyThrHh1sAlaPheHh1sAlaAlaCybaPheCy 400
 1317 TGGTCAATCCCGTGGCTCATGGAATCAATGATTTACAGCGCTGCTGCTGCTGCTGCT 1376
 401 AlaThrgInlybValGlyGluGlnbencybIleLybLeuIlePheGlnGlyProIleAsp 420
 1377 GCTACACACACTGTGGGAGAGCAAACTGCATCAAAATTAATTTCCAAAGTCCAAATTGC 1436

RESULT 5
 AX686405
 LOCUS AX686405 5921 bp DNA linear PAT 29-MAR-2003
 DEPOSITION Sequence 3 from Patent WO02059611.
 ACCESSION AX686405
 VERSION AX686405.1 GI:29372143
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 Powers, S., Mu, D., Xiang, P. and Peng, Y.
 1 Diagnosis and treatment of cancer using mammalian pellino polypeptides and polynucleotides
 JOURNAL Patent: WO 02059611-A 3 01-AUG-2002;
 Tularik Inc (US)
 FEATURES
 source Location/Qualifiers
 1..5921
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"

CDS
/db_xref="taxon:9606"
501...1763
/note="unnamed protein product; human pellino 2"
/codon_start=1
/protein_id="CAD86576.1"
/db_xref="GI:2937214"

/translation="MPSQGEHCAENKPEVKYGEIVLVGYNGALPNGDRGRKRFALYKRPKANGKSTVHISTPOASKAISCGQHSISYLSKQVTVVETHDKDTDMFQVGRSTESPIDFVITDTISGQNTDEAOITOSTISRFACRIVCDNEPYTARIFAAPDSSKNIPLGEKAAKKKPNPGHMDGTTNNVLMHPRGFTESQGWREISVCCDVTRETRSAQQRGLVBSSTNVLDQSLDLCATLLMTRADGLFHTPOKHEALRQEIINAARQPCVGNLTLPSPINKEVEEKQPAWYLSCGHNVHMYHMSDTEANBRECPMCRVGPYPLMVLGCEAGFYVDAGPPTPAFTPCGHVCSERKATWSQIPPHGTTHAFHACPCFATQVGEONCIKLFQGPID"

ORIGIN

Alignment Scores:

Pred. No.: 1,93e-176 Length: 5921
Score: 2290.00 Matches: 420
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-041-030-4 (1-420) x AK686405 (1-5921)

QY 1 MetPheSerProGlyGlnGlnGlnHisCysAlaProAsnLysGluProValLysTyrGly 20
DB 501 ATGTTTCCCTGGCGAGGAGGAAACACTGCGCCCAATAGAGAGCAGTAAATACGGG 560
QY 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgLys 40
DB 561 GAGCGTGTGTCTCGGTACAAATGTTGCTTTACCAATCGAATGAGAGACGAGGAAA 620
QY 41 SerArgPheAlaLeuTyrTyrAsnGlyProLysAlaAsnGlyValLysProSerThrValHis 60
DB 621 AGTAGATTGTCCTTCAACAGCGGCCCAAGGCAAAAGTGTCMAACCCAGCACCGTCCAT 680
QY 61 ValLysSerThrProGlnLysSerLysAlaLysSerCysValGlyGlnHisSerLysSer 80
DB 681 GTGATATCCAGCCCGCCAGCATTCAGGCTATTCAGTGCAGAAAGGTCAACACAGTATATCC 740
QY 81 TyrThrLeuSerArgAsnGlnThrValValAlaGluTyrThrHisAspLysAspThrAsp 100
DB 741 TACACTTTGTCAAGAAATCAGACTGCTGCTGAGAGTACACATGATTAAGATACGAT 800
QY 101 MetPheGlnValGlyArgSerThrGlySerProLysAspPheValValThrAspThrIle 120
DB 801 ATGTTTCAGGTGGCGAGATCAACAGAAACCCCTATCGACTTCTGTTGTCAACAGACGATT 860
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrLysSerArgPheAla 140
DB 861 TCTGGAGCCCAACACCGAAGCCAGATCACACAGACCCATATCCAGTTCCGCC 920
QY 141 CysArgLysValCysAspArgAsnGluProTyrThrAlaArgLysPheAlaAlaGlyPhe 160
DB 921 TCGAGGATCGTGGCGACAGAAATGAACCTTACACAGCACGGAATATCCCGCGGATTT 980
QY 161 AspSerSerLysAlaAsnLysPheLeuGlyGlyLysAlaAlaLysTyrLysAsnProAspGly 180
DB 981 GACTCTTCCAAAACATATTTCTTGGAGAAAGCGCAAGTGGAAAAACCCCGACGCGC 1040
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLysValMetHisProArgGlyGlyPheThr 200
DB 1041 CACATGATGGGCTCATATATGGCGTCTCGGTATGATCATCCACAGAGGGGCTTCACC 1100
QY 201 GluGluSerGlnProGlyValTyrParGluLysSerValCysGlyAspValTyrThrLeu 220
DB 1101 GAGAGATCCCAAGCCCGGGGTCTGGCGAGATCTCTGCTGAGAGATGTGTACACCTTG 1160
QY 221 ArgGluThrArgSerAlaGlnGlnArgLysLysLeuValGluSerGluThrAsnValLeu 240
DB 1161 CGAGAAACGAGGTGGCCGACGAAACGAGAAAGCTGTGTGAAAAGTGAAGACCAACGTCTG 1220

QY 241 GlnAspGlySerLeuLysAspLeuGlyValAlaThrLeuLeuTyrParGThrAlaAspGly 260
DB 1221 CAGGAGGCTCCCTCATTTGACCTGTGTGGGCCACTCTCTCTGGAGAAACAGCATGGG 1280
QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnLysAsnAlaAla 280
DB 1281 CTTTTCATCTCCCACTCAGAAAGCATAGAAAGCCCTCCGGCAGAGATTAACCCGCC 1340
QY 281 ArgProGlnGlyProValGlyLeuAsnThrLeuAlaPheProSerLysAspArgLysGlu 300
DB 1341 CGGCTTCAGTGTCTCTGGGCTCAACACCTGCGCTTCCACGATCAACAGGAAAG 1400
QY 301 ValValGluGluLysGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
DB 1401 GTGGTGGAGAGAACAGACCCCTGGGCAATCTCAGTTGTGGCCAGTGCAGGATACAC 1460
QY 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
DB 1461 AACTGGGCGCATCGAGTGAACGAGGCCCAACGAGAGGAGTGTCCCATGTGCAGACT 1520
QY 341 ValGlyProTyrTyrAlaProLeuThrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
DB 1521 GTGGGCCCCCTATGTGCTCTCTGCTGTGAGGAGGAGATTTATGTAGACGACAGA 1580
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
DB 1581 CGGCCAATCATGCTTTTCACTCCCTGTGACACGCTGTCTGGAGAAATGTGCAAAATAC 1640
QY 381 TrpSerGlnLysProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
DB 1641 TGTCTCAGATCCCTTCTCTCATATGAACTATGATTTTCAACGCTGTGCTTCTGT 1700
QY 401 AlaThrGlnLeuValGlyGlnGlnAsnCysAlaLysLeuLysLeuLysPheGlnGlyProLysAsp 420
DB 1701 GCTACACAGCTGTGTGGGAGCAAACTGCATCAATTAATTTTCAAGGTCAATGAC 1760

RESULT 6
AR483568 1263 bp DNA linear PAT 14-MAY-2004
LOCUS AR483568
DEFINITION Sequence 7 from patent US 6703487.
ACCESSION AR483568
VERSION AR483568.1 GI:4724652
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1263)
AUTHORS Bird,T.A. and Coeman,D.J.
TITLE Human pellino polypeptides
JOURNAL Patent: US 6703487-A 7 09-MAR-2004;
FEATURES
location/Qualifiers
1..1263
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,11e-176 Length: 1263
Score: 2283.00 Matches: 419
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 99.69% Indels: 0
DB: 6 Gaps: 0

US-10-041-030-4 (1-420) x AR483568 (1-1263)

QY 1 MetPheSerProGlyGlnGlnGlnHisCysAlaProAsnLysGluProValLysTyrGly 20
DB 1 ATGTTTCCCTGGCGAGGAGGAAACACTGCGCCCAATAGAGAGCAGTAAATACGGG 60
QY 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgLys 40

Db 61 GACCTGCTGCTCGGCTCAATGCTCTTACCCATGAGATGAGGACGAGGAAA 120
 Qy 41 SerArgPheAlaLeuTyrLybArpProLybAlaAsnGlyValLybProSerThrValHis 60
 Db 121 AGTAGATTGGCTCTCAACAAGCGGCCCAAGGCAAAATGTGTCAAAACCGACGACGTCAT 180
 Qy 61 ValIleSerThrProGlnAlaSerLybAlaIleSerCylbLybGlyGlnHisSerIleSer 80
 Db 181 GTGATATCCAGCGCCCAAGGCGATCCAAAGGCTATCAGCTGCGCAAGAGTCAACCACTATATATCC 240
 Qy 81 TyrThrLeuSerArgbAngInThrValValValGlyTyrThrHisbApLybApbThrAp 100
 Db 241 TACACTTTGTCAAGAAATCAGACTGTGTGTGTGAGTACACATGATTAAGATTACGAT 300
 Qy 101 MetPheGlnValGlyArgSerThrArgLybSerProIleAspPheValIleThrAspThrIle 120
 Db 301 ATGTTTCAGGTGGGCGATCAACAGAAAGCCCTATGACTCTGTTGTCAACGACGACGATT 360
 Qy 121 SerGlySerGlnAsnThrAspGlyValGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 361 TCTGGCAGCCAGAAACAGCGAGCCAGATCAACAGAGCCCATATCCAGGTTCCGCC 420
 Qy 141 CysArgIleValCysbApArgbAngInProTyrThrAlaArgIlePheAlaIleGlyPhe 160
 Db 421 TGCAGATCGTGTGCGACAGAAATGAACCTTACACAGCAGAAATTCGCCCGGAGATT 480
 Qy 161 AspSerSerLybAsnIlePheLeuGlyGlyValAlaAlaIleTyrLybAsnProAspGly 180
 Db 481 GACTCTTCCAAAACATATTTCTTGAGTAAAGCGAACGAAAGTGGAAAAACCCGACGCG 540
 Qy 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 Db 541 CACATGATGGGCTCACTACTAATAGGCGCTGTGTGTGATGATCACAAGGCGGCTTCACC 600
 Qy 201 GluGlnSerGlnProGlyValIleTyrArgGlyIleSerValCysGlyAspValIleThrLeu 220
 Db 601 GAGAGATCCAGCGCGGCTGTGCGAGATCTGTGTGTGAGATGTGTACCTTGG 660
 Qy 221 ArgGlyThrArgSerAlaGlnIleArgGlyLybLeuValGlySerGlyThrAsnValLeu 240
 Db 661 CGAGAAACAGGCTCGGCCCAAGCAAGAAAGCTGTGGAAGTGAACCAACGTCCTG 720
 Qy 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrArgThrAlaAspGly 260
 Db 721 CAGGACGGCTCCCTCATTTGACTGTGTGTGGGCCACTCTCTGTGAGAAACAGCAGATGG 780
 Qy 261 LeuPheHisThrProThrGlnLybHisIleGlyAlaLeuArgGlnGlyIleAsnAlaAla 280
 Db 781 CTTTTTCATATCTCACTCAAGAACCATAGAAAGCCCTCGGAGGAGATTAAACGCGCC 840
 Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLybGly 300
 Db 841 CGGCTCATAGTGTCTGTGGGGCTCAACACCTGCGCTTCCCGACGATCAACAGAAAGAG 900
 Qy 301 ValValGlyGlyGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
 Db 901 GTGTGTGAGAGAGAGCAGCCCTGAGGCTATCTCAGTTGTGGCCAGCGATATACAC 960
 Qy 321 AsnTyrGlyHisArgSerArgbThrGlyAlaAsnGlyValGlyCysProMetCysbArgThr 340
 Db 961 AACTGGGGCATGAGAGTGAACGAGAGCCAAAGAGGAGTGTCCCATGTCCAGAGACT 1020
 Qy 341 ValGlyProTyrValProLeuTyrPheGlyCysGlyValGlyPheTyrValAlaAlaGly 360
 Db 1021 GTGGGCCCCATATGTGCTCTGTGCTGTGTGAGCAGAGATTTTATGAGACGAGGA 1080
 Qy 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlyLybSerAlaIleTyr 380
 Db 1081 CCGCAACTCATGTCTTCACTCCCTGTGAGACGTGTGTGCGAGAGTGTGCAAAATATAC 1140
 Qy 381 TrpSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
 Db 1141 TGTCTCAGATCCCGTGTGCTCATGAACTCATGCAATTTCACGCTGCTTCCCTTCTGT 1200

Qy 401 AlaThrGlnLeuValGlyGlyGlnAlaPheCysAllelybLeuIlePheGlnGlyProIleAsp 420
 Db 1201 GCTACACAGCTGTGTGGGAGCAAAATGTCATCAAAATTAATTTTCCAAAGTCCAAATGAC 1260
 RESULT 7
 AX298079 1263 bp DNA linear PAT 26-NOV-2001
 LOCUS Sequence 7 from Patent WO0183739.
 DEFINITION AX298079
 ACCESSION AX298079
 VERSION AX298079.1 GI:17128160
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 Bird, T. A. and Cosman, D. J.
 Human peptidic polypeptides
 Patent: WO 0183739-A 7 08-NOV-2001;
 JOURNAL IMMUNEX CORPORATION (US)
 FEATURES
 source 1..1263
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,116-176 Length: 1263
 Score: 2283.00 Matches: 419
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 1
 Query Match: 99.69% Indels: 0
 DB: 6 Gaps: 0
 US-10-041-030-4 (1-420) x AX298079 (1-1263)
 Qy 1 MetPheSerProGlyGlnGlyGlnHisCysAlaProAsnLybGlyProValIleTyrGly 20
 Db 1 ATGTTTCCCTCGGCCAGAGGAAACATGCGGCCCAATAGAGGCACTGAAATACGCG 60
 Qy 21 GluLeuValValLeuGlyTyrbAnglyAlaLeuProbAnglyAspArgGlyArgbGlyb 40
 Db 61 GAGCTGTGTGTCTCGGCTCAATGATGTGTCTTACCAATGAGATAGAGAGCGAGGAAA 120
 Qy 41 SerArgPheAlaLeuTyrLybArpProLybAlaAsnGlyValLybProSerThrValHis 60
 Db 121 AGTAGATTGGCTCTCAACAAGCGGCCCAAGGCAAAATGTGTCAAAACCGACGTCAT 180
 Qy 61 ValIleSerThrProGlnAlaSerLybAlaIleSerCylbLybGlyGlnHisSerIleSer 80
 Db 181 GTGATATCCAGCGCCCAAGGCGATCCAAAGGCTATCAGCTGCGCAAGAGTCAACCACTATATCC 240
 Qy 81 TyrThrLeuSerArgbAngInThrValValValGlyTyrThrHisbApLybApbThrAp 100
 Db 241 TACACTTTGTCAAGAAATCAGACTGTGTGTGTGAGTACACATGATTAAGATTACGAT 300
 Qy 101 MetPheGlnValGlyArgSerThrArgLybSerProIleAspPheValIleThrAspThrIle 120
 Db 301 ATGTTTCAGGTGGGCGATCAACAGAAAGCCCTATGACTCTGTTGTCAACGACGACGATT 360
 Qy 121 SerGlySerGlnAsnThrAspGlyValGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 361 TCTGGCAGCCAGAAACAGCGAGCCAGATCAACAGAGCCCATATCCAGGTTCCGCC 420
 Qy 141 CysArgIleValCysbApArgbAngInProTyrThrAlaArgIlePheAlaIleGlyPhe 160
 Db 421 TGCAGATCGTGTGCGACAGAAATGAACCTTACACAGCAGAAATTCGCCCGGAGATT 480
 Qy 161 AspSerSerLybAsnIlePheLeuGlyGlyValAlaAlaIleTyrLybAsnProAspGly 180
 Db 481 GACTCTTCCAAAACATATTTCTTGAGTAAAGCGAACGAAAGTGGAAAAACCCGACGCG 540

QY	181	HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProAspGlyValGlyPheThr	200
Db	541	CACATGATGAGGACCTCACTAAATGGCGCTCGGTGGATGATCCACGAGGGGGCTTCACC	600
QY	201	GlulGluSerGlnProGlyValTTPArgGluIleSerValCysGlyAspValTyrThrLeu	220
Db	601	GAGAGATCCACGCCCCGGGGTCTGGGCGGAGATCTCTGTCTGTGGAGATGTTCACCTTG	660
QY	221	ArgGluThrTrpArgSerAlaGlnGlnAlaArgGlyValSerValGluSerGluThrAsnValLeu	240
Db	661	CGAGAAACACAGTCTGGCCCCACCAACAGAGAAAGCTGGGAGAAAGTGAACCAACGATCTCTG	720
QY	241	GlnAspGlySerLeuIleAspLeuCysGlyValAlaThrLeuLeuThrArgThrAlaAspGly	260
Db	721	CAGACGGCTCCCTCATTTGACTCTGTGTGGGGCCACTCTCTCTGGAGAACGACGATGGG	780
QY	261	LeuPheHisThrProThrGlnValHisIleGluValaLeuArgGlnGluIleAsnAlaAla	280
Db	781	CTTTTTCATATCTCCAACTCAGAAAGCACATAGAAAGCCCTCCGGCAGAGATTAACGCCGCC	840
QY	281	ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgGlyGlu	300
Db	841	CGGCTCAGTGTCTGTGGGGCTCAACACCTGGGCTTCCCGAGATCAACAGAGAAAGAG	900
QY	301	ValValGluGluIlyGlnProTTPAlaTyrLeuSerCysGlyHisValHisGlyTyrHis	320
Db	901	GTGGTGGAGAGAAAGCACCCTCGGSCATATCTCAAGTTGTGGCCACGTGCACGGTCACAC	960
QY	321	AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr	340
Db	961	AACGGGGGCGATCGGAGTGACACGGAGGCGCAACGAGAGAGTGTCCATGTGAGAGACT	1020
QY	341	ValGlyProTyrValProLeuThrLeuGlyCysGluValGlyPheTyrValAspAlaGly	360
Db	1021	GTGGGCCCCATMGTCCTCTGTGGCTTGAGCTGTGGAGGAGATTTATGTAGACGCMGA	1080
QY	361	ProProThrHisAlaPheThrProCysGlyHisValCysSerGluIlySerAlaIlyTyr	380
Db	1081	CGGCAACTCATGCTTTCACCTCCCTGTGGACACGHTGCTCGGAGAAAGTCTGCAAAATAC	1140
QY	381	TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys	400
Db	1141	TGGTTCAGATCCCGCTTCCCTCATGGAACCTCATGATTCACGCTGTGGCCCTTTCGT	1200
QY	401	AlaThrGlnLeuValGlyGluAsnCysIleLeuValIlePheGlnGlyProIleAsp	420
Db	1201	GCTACACAGCTGGTGGGGAGCAAACTGCATCAATTAATTTTCCAAAGTCCAAATTGAC	1260
RESULT 8			
AR483567			
LOCUS	AR483567	1260 bp	DNA linear PAT 14-MAY-2004
DEFINITION	Sequence 5 from patent US 6703487.		
ACCESSION	AR483567		
VERSION	AR483567.1 GI:47246651		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1260)		
AUTHORS	Bird, T.A. and Cosman, D.J.		
TITLE	Human pellino polypeptides		
JOURNAL	Patent: US 6703487-A 5 09-MAR-2004;		
FEATURES	Location/Qualifiers		
source	1..1260		

ORIGIN /mot_cype--genomic DNA

Alignment Scores:

Pred. No.:	6.44e-168	Length:	1260
Score:	2175.50	Matches:	399
Percent Similarity:	97.38%	Conservative:	10

Best Local Similarity:	95.00%	Mismatches:	10
Query Match:	95.00%	Indels:	1
DB:	6	Gaps:	1

US-10-041-030-4 (1-420) X AR483567 (1-1260)

QY	1	MetPheSerProGlyGlnGlnGlnHisCysAlaAspProLeuValProValIleTyrGly	20
Db	1	ATGTTTTCCTCCGGGCGCAGAGAAACCCAGCGCCCCCAACAGAGCGGTAAATACCGG	60
QY	21	GluLeuValValIleGlnGlyTyrAsnGlyValAlaLeuProAsnGlyAspArgGlyArgArgIys	40
Db	61	GAGCTGGTGGTCTCGGGGTACAATGGTGTCTTACTTAATGGTGAACGGGCGCAGAGGAAA	120
QY	41	SerArgPheAlaLeuTyrTyrAsArgProLeuValAsnGlyValIysProSerThrValHis	60
Db	121	AGCAGATTTCCTCTATTAAGCGGACCTACCGCAGTGTGTCAAAACCCAGACCAATCCAC	180
QY	61	ValIleSerThrProGlnAlaSerTyrAlaIleSerCysIysGlyGlnHisSerIleSer	80
Db	181	ATGGCTTCACACACAGCGCTCCAGGCGCATTCACGTCACAGAGACATCACAGCATATCG	240
QY	81	TyrThrIleuSerArgAsnGlnThrValValValGlnTyrThrHisAspLysAspThrAsp	100
Db	241	TACACGTTGTACGGAGCGCAGACGGTGTGGTGAAGTACACACAGATTAAGACACGGAC	300
QY	101	MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle	120
Db	301	ATGTTTCAGGTGGGCGAGGTCAACAGAAAGCCCATTTGCTTCGGTGCACAGACGGTT	360
QY	121	SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla	140
Db	361	TCCGGCGGTCAAGACGAAGAT--GCCAGATCCACAGAGCACCATCTTGAGTTGCCA	417
QY	141	CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe	160
Db	418	TGCAGAGATCGTGTGACAGGACGAGCCATATACAGACGATATTCGGCGGAGAGATTTC	477
QY	161	AspSerSerIysAsnIlePheLeuGlyGlnValAlaAlaIysTrpLysAsnProAspGly	180
Db	478	GATTCCTTCCAAAATATCTTTCTTGGAGAGAAACAGCAAAATGAAAACCTCGATGGA	537
QY	181	HisMetAspGlyLeuThrThrAsnGlyValIleuValMetHisProArgGlyGlyPheThr	200
Db	538	CACATGATGATGACTCACATCAACATGGTGTCTTACTGATGACACCGCAAGAGGCTTCC	597
QY	201	GluGluSerGlnProGlyValTyrArgGluIleSerValCysGlyAspValTyrThrIleu	220
Db	598	GAGGAATCCAGCGCTGAGTCTGAGAGAGAGATCTCTGCTGTGGGGATGTGTACACCTTG	657
QY	221	ArgGluThrArgSerAlaGlnGlnArgGlyIysIleuValGluSerGluThrAsnValIleu	240
Db	658	CGAAGACACAGAGTGGCCACGACGAGGGGAAAGCTGGGAAAGTGAACCCAGTCTTG	717
QY	241	GlnAspGlySerLeuIleAspLeuCysGlyValIleThrLeuLeuTyrArgThrAlaAspGly	260
Db	718	CAAGACGGCTCCCTCATTTGACTGTGTGGGGCCACTCTCTCTGGAGAACCGCAGATGGC	777
QY	261	LeuPheHisThrProThrGlnLysHisIleGluValAlaLeuArgGlnGluIleAsnAlaIle	280
Db	778	CTTTTTCACGCTCTACTACAGAACACATAGAACCTCCCGCAGAGATCAATGACGCC	837
QY	281	ArgProGlnCysProValGlyLeuAsnThrIleuAlaPheProSerIleAsnArgLysGlu	300
Db	838	CGACCCCACTGCCCGTGGGCTTAAACACCTGGACCTTCCCAACATCAACCGAAGGAA	897
QY	301	ValValGluGluLysGlnProTrpAlaTyrIleuSerCysGlyHisValHisGlyTyrHis	320
Db	898	GTTGGTGAAGGAAGACAGCCCTGGGCATCTGAGCTGGCGGCATGTGCAAGGCTACAC	957
QY	321	AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgLysCysProMetCysArgThr	340
Db	958	AGCTGGGGCTCATCGAGAGACGCGGAAGCCCAACAGAGGGAGTGTCCATGTGCAGACT	1017

QY 341 ValGlyProGlyValPLeuTTPLeuGlyCyseGluAaglYpHeTYValAaPAlaGly 360
 Db 1018 GTGGGCCCCAAGCCCTCTCTGCTGGCTGGTGAAGCAGATTTTATGTCGATGCGGGA 1077
 QY 361 ProProthHhAAlaPheThrProCyGlyVhIValCyseSerGlyuysSerAlaVtYr 380
 Db 1078 CCCCCAATCAAGCTTTCACCCCTGCGGGACGTGTTCAGAAAGTCTGCCAAGTAC 1137
 QY 381 TPSeSerGlnIleProLeuProHhAglYThrHhAAlaPheHhAAlaAaCyseProPheCy 400
 Db 1138 TGGTCGAGATCCCACTGCCCCACGGAACGACGCTTCATGCGCTGCTGCTGCTG 1197
 QY 401 AlaThGlnLeuValGlyGlnGlnAerCyseIleValLeuIlePheGlnGlyProIleAaP 420
 Db 1198 GCCACGACGTGGTGGTGAACGAAGTCATCAATGATTTTCCAAAGGTCCAGTGAC 1257

RESULT 9
 AX298077 1260 bp DNA linear PAT 26-NOV-2001
 LOCUS Sequence 5 from Patent WO0183739.
 AX298077
 AX298077.1 GI:17128159
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 Bird, T.A. and Cosman, D.J.
 Human peilin polypeptides
 Patent: WO 0183739-A 5 08-NOV-2001;
 IMMUNEX CORPORATION (US)

FEATURES
 source
 1..1260
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10090"

ORIGIN

Alignment Scores:
 Pred. No.: 6 44e-168 length: 1260
 Score: 2175.50 Matches: 399
 Percent Similarity: 97.38% Conservative: 10
 Best Local Similarity: 95.00% Mismatches: 1
 Query Match: 95.00% Gaps: 1
 DB: 6

US-10-041-030-4 (1-420) x AX298077 (1-1260)

QY 121 SerGlySerGlnAaThrAaPAlaAaglIleThGlnSerThrIleSerAaPheAla 140
 Db 361 TCCGGGCGTCAGAACGAAGAT---GCCAGATCACACAGACCAATCTCTAGTTCGCA 417
 QY 141 CyAArgIleValCyseAaPArgAaGlnPProTYThrAlaArgIlePheAlaAaglYpHe 160
 Db 418 TGCAGATGCTGTGTACAGAAACGACCAATATACAGACGCAATATTCGGGCGAGATTC 477
 QY 161 AaPSeSerGlyAaAaIlePheLeuGlyGlnIleValAAlaVtYrTPYsaAaPProAaP 180
 Db 478 GATTCCTCCAAATAATCTTCTTGGAGAAAGACGCAAAATGAAAAACCTGATGGA 537
 QY 181 HhMetAaPGLYLeuThrThrAaGlnValLeuValMetHhPProAaArgIlyGlyPHeThr 200
 Db 538 CACATGATGAGACTCACTACATCAATGATGCTGATGATGATCACCACCGAGAGCTTCACC 597
 QY 201 GlnGlnSerGlnProGlyValTPAArgGlnIleSerValCyseGlyAaPValTYrThrLeu 220
 Db 598 GAGGAATCCAGGCTGAGCTGAGAGAGATCTGTCTGTGGGATGTATACACCTTG 657
 QY 221 ArgGlnThrAaSerAaAaglGlnAaArgIlyLeuValGlySerGlnThrAaAaValleu 240
 Db 658 CGAGACACAGAGTCGCGCCACGAGAGGGAAGCTGTGGAAGTACACACCTCTG 717
 QY 241 GlnAaPGLYSerIleAaPLeuCyseGlyAAlaThrLeuLeuTPAArgThraAaPGLY 260
 Db 718 CAAGACGCTCTCTCAATGACCTGTGTGGGCGACCTCTCTGGAAGAACCGCAGATGGC 777
 QY 261 LeuPheHhAaThrProThrGlnIleVhIleGlnAlaLeuAaGlnIleValAaAaIleAa 280
 Db 778 CTTTTCAGGCTCTCACTCAAGAGCAGATGAAGCCCTCGGAGGAGATCAATGACGCT 837
 QY 281 ArgProGlnCyseProValGlyLeuAaThrLeuAAlaPheProSerIleAaAaArgIlyGln 300
 Db 838 CGACCCCAAGTCCGCGTGGGCTTACACCTGCGCTTCCACGATCAACGGAAGAA 897
 QY 301 ValValGlnGlnIleProThraAlyrLeuSerCyseGlyVhIleVhIleGlyTYrHhA 320
 Db 898 GTGGTGAAGAGAAAGACGCTGGGCTATCACTGAGCTGCGGCATGTGACGCTACAC 957
 QY 321 AaAaThrGlnIleAaSerAaThrGlnAAlaAaGlnAaArgGlnCyseProMetCyseAaGln 340
 Db 958 AGCTGGGCGCATGAGGAGCAACGCGAAGCAACGAGAGAGTCCATGTGACAGACT 1017
 QY 341 ValGlyProTYrValPLeuTTPLeuGlyCyseGluAaglYpHeTYValAaPAlaGly 360
 Db 1018 GTGGGCCCCAAGCCCTCTCTGCTGGCTGGTGAAGCAGATTTTATGTCGATGCGGGA 1077
 QY 361 ProProthHhAAlaPheThrProCyGlyVhIleValCyseSerGlyuysSerAlaVtYr 380
 Db 1078 CCCCCAATCAAGCTTTCACCCCTGCGGGACGTGTTCAGAAAGTCTGCCAAGTAC 1137
 QY 381 TPSeSerGlnIleProLeuProHhAglYThrHhAAlaPheHhAAlaAaCyseProPheCy 400
 Db 1138 TGGTCGAGATCCCACTGCCCCACGGAACGACGCTTCATGCGCTGCTGCTGCTG 1197
 QY 401 AlaThGlnLeuValGlyGlnGlnAerCyseIleValLeuIlePheGlnGlyProIleAaP 420
 Db 1198 GCCACGACGTGGTGGTGAACGAAGTCATCAATGATTTTCCAAAGGTCCAGTGAC 1257

RESULT 10
 BC072891 1646 bp mRNA linear VRT 03-AUG-2004
 LOCUS Xenopus laevis MGC80329 protein, mRNA (cDNA clone MGC:80329
 IMAGE:5074048), complete cds.
 BC072891
 BC072891.1 GI:49119163
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Xenopus laevis (African clawed frog)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

Xenopodinae; Xenopus; Xenopus.

REFERENCE 1 (bases 1 to 1646)
AUTHORS Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P.
TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus Initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 12454917
AUTHORS

2 (bases 1 to 1646)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heich,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stappleton,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Ueda,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loggiano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Pahey,J., Helton,E., Kettelman,M., Madan,A., Rodigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E., Scherch,A., Schein,T.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
AUTHORS Klein,S. and Gerhardt,D.S.

3 (bases 1 to 1646)
Direct Submission
TITLE Submitted (07-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
JOURNAL NIH-WGC Project
REMARK Contact: XGC help desk
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Igor Dawid
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@cgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Ohi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liles Prabu, Parvaneh Saeedi, J.R Santos, Angeliqne Scherch, Ursula Skalska, Duane Smallus, Jeff Stott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRK Plate: 152 Row: h Column: 6
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.
Location/Qualifiers
1.1646
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:80329 IMAGE:5074048"
/tissue_type="Ovary, Xenopus"
/clone_id="NICHD_XGC_Ov1"
/lab_host="DH10B"
/note="Vector: PCMV-SPOrt6"
1.1646
gene

/gene="MGC80329"
/db_xref="locusID:443954"
336.1595
/gene="MGC80329"
/product="MGC80329 protein"
/protein_id="AAH72891.1"
/db_xref="GI:4919164"
/db_xref="locusID:443954"
translation="MFSPSQEBHCAPSKPEVYGEVLVVLGNGCLPNDRGRKSRPA
LFRKRSKGVNRSVHVISTQASAKISRGQHSISFTLSQTVVAYTHDKQDMF
QIGRTSEPIDFVVDIVSGNQNDTQVTSQTSISFACRIVCDNPPTATFAAGD
SKNKFLDEKAKAMPDGHGDTLNQVWHPKGGTDSRPMVEISVCGVYT
LRETSADRGKLVENETNIQDGSIVLDCATILMTARDGLHTPTQKHEALRQEI
NNAAPQCVGANTLAFPSINKDVEDQPMAYLKCGVYHGHMGRHSDEANREK
PMKRTIGYVPLMLCEAGFYVDGPPTHAFPCGHVCSKSAKTSQIPLPHGTHAF
HAACFCAIQLTGESSCVKLFGQVVD"

ORIGIN

Alignment Scores:

Pred. No.:	9,24e-165	Length:	1646
Score:	2138.50	Matches:	383
Percent Similarity:	96.90%	Conservative:	24
Best Local Similarity:	91.19%	Mismatches:	12
Query Match:	93.38%	Indels:	1
DB:	5	Gaps:	1

US-10-041-030-4 (1-420) x BC072891 (1-1646)

QY 1 MecPheSerProGlyGlnGluGluHscYsAlaProAenLyGluProValIleTyrGly 20
|||||
DB 336 ATGTTTCCCGACGACGAGGAGACATTCGCCGCCCATGAAAGACCCGTAAATACGGA 395
|||||

QY 21 GluLeuValIleGluGlyTyrAenGlyAlaLeuProAenGlyAspArgGlyAspArgGly 40
|||||
DB 396 GAGCTGGTGCTGCTGGGAGACAAACGCTGTCTCCAAATGGGGATCGAGAGAGCGGAAA 455
|||||

QY 41 SerArgPheAlaLeuTyrIleAspProLyAlaAenGlyValIleProSerThrValHis 60
|||||
DB 456 AGCAGATTTCACCTTTCAAAGCCCAATATTAACGGGGTGAAGCCAGCCAGCTTTCAT 515
|||||

QY 61 ValIleSerThrProGlnAlaSerLyAlaIleSerCybLyGlyGlnHisSerIleSer 80
|||||
DB 516 GTGATTTCACCCGACGAGCTTCGAAGCAATCAGCAGAGAGGCGCAACAGCATCTCA 575
|||||

QY 81 TyrThrLeuSerArgAsnGlnThrValValValGluTyrThrHisAspIleAspThrAsp 100
|||||
DB 576 TTCACATTATCCAGAAAGTCAGACTGTGTGGTAATCACCCATGACAGACACAGAT 635
|||||

QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120
|||||
DB 636 ATGTTTCAGATTGGAAAGTCACAGAAAGCCCATTCGTTTGTAATCTATACGGTT 695
|||||

QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
|||||
DB 696 TCTGGCAACCAAGAAC---GATGAACACAGATCACAAAGTACATTTCAAGATTGCA 752
|||||

QY 141 CysAaGllleValCysAspArgAsnGluProTyrThrAlaAglIlePheAlaIleAlaIlePhe 160
|||||
DB 753 TGTAGAAATAGTGTGGATAGAAACCCGCGTACACAGGTAGAAATTTGCTGCGACAGATT 812
|||||

QY 161 AppSerSerLyAsnIlePheLeuGlyGluLyAsnAlaIleTyrLyAsnProAspGly 180
|||||
DB 813 GACTCTCCAAAATATATTTCTTGTGTAGAAAGCGCCCAAGTGAAGAAATCTTCACGCT 872
|||||

QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisAspArgGlyGlyPheThr 200
|||||
DB 873 CATATGACACGACCTAATACCAATGGGTTCTGTGATGATCAACAAAGGAGGAGTTTACT 932
|||||

QY 201 GluGluSerGlnProGlyValIleTyrArgGluIleSerValCysGlyAspValIleThrLeu 220
|||||
DB 933 GATGAATCAGCGCTGTGTGTGTGGGAGAAATTCGTGCTGTGAGAGATGTTATACATTA 992
|||||

QY 221 ArgGluThrArgSerAlaGlnGlnAlaGlyLeuValGluSerGluThrAlaValLeu 240
 DB 993 AGAGAAACAGATCTCTCTCAACAAAGGCAACCTTGTGAAATGAACCAATATCTT 1052
 QY 241 GlnAerGlySerLeuIleAerLeuCyGgiAlaIleThrLeuLeuTrpArgThraAlaerGly 260
 DB 1053 CAGGATGGTCTCTCTTGTGACTTATGTGGGCCACCTCTTGGGAAACAGCGATGGG 1112
 QY 261 LeuPheHisThrProThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 DB 1113 TTGCTTCATACCGCGCTCAAAAACATTTGAAGCTTTGAGACAGAAATTAATGTCTGC 1172
 QY 281 ATGProGlnCyProValGlyLeuAerThrLeuAlaPheProSerIleAerAerGlyGln 300
 DB 1173 AGGCTCACTGCCCCCTTGGATTAAACCTTTTCCAGTATCAACGCAAGAT 1232
 QY 301 ValValGluGluGlnProTrpAlaIleLeuSerCyGgiAlaIleValHisGlyIleHis 320
 DB 1233 GTTGTGAAGACAAACAGGCTTGGCTTACCTCAAGTGTGATGATGCTATCAT 1292
 QY 321 AarTrpGlyHisAerSerAerThrGluAlaAerGlnAerGluCyProMetCyAerGln 340
 DB 1293 AACTGGGACATCGCAGTACACAGAACCAACGACGAGTGTCCATGTGTAGACA 1352
 QY 341 ValGlyProTrpValProLeuTrpLeuGlyCyGgiAlaGlyPheTrpValAerAlaGly 360
 DB 1353 ATTTGCTTCATACCGCGCTCAAAAACATTTGAAGCTTTGAGACAGAAATTAATGTCTGC 1412
 QY 361 ProProThrHisAlaPheThrProCyGgiAlaValCySerGlnIleSerAlaIleVal 380
 DB 1413 CCGCTCACTCAACGCTTCAAGGCTTGGACAGCTGTGCTGCAAAAATCTGCAAAAT 1472
 QY 381 TrpSerGlnIleProLeuProHisGlyIleHisAlaPheHisAlaAerCyProPheCy 400
 DB 1473 TGTGTCCCAATTCCTGCTCCACATGTGTACACGCTTTCACGCTTGTCTCTGT 1532
 QY 401 AlaThrGlnLeuValGlyGluGlnAerCyGgiAlaValLeuIlePheGlnGlyProIleAer 420
 DB 1533 GCATATCAATTCATCTGAGGAGAGAGCTGTGCAAACTTATTTTCCAGGACCTGTTGAC 1592
 RESULT 11
 BC063200 1936 bp mRNA linear VRT 30-JUN-2004
 LOCUS Xenopus tropicalis hypothetical protein MG75650, mRNA (cDNA clone
 DEFINITION MG75650 IMAGE:5379633), complete cds.
 ACCESSION BC063200
 VERSION BC063200.1 GI:38649365
 KEYWORDS MGC.
 SOURCE Xenopus tropicalis (Silurana tropicalis)
 ORGANISM Xenopus tropicalis
 Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 1936)
 Struhsberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buec, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stalderon, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L.,
 Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carinaz, P., Frange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, J.C., Steichen, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length

JOURNAL
 PUBMED
 12477932
 2 (bases 1 to 1936)
 REFERENCE
 Klein, S. and Struhsberg, R.
 Direct Submission
 Submitted (02-DEC-2003) National Institutes of Health, Xenopus Gene
 Collection (XGC), National Institute of Child Health and Human
 Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
 20892-7510, USA
 NIH-MGC Project
 CONTACT: XGC help desk
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: Robert M. Grainger
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bgsc.bc.ca

REMARK
 COMMENT
 human and mouse cDNA sequences.
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1936)
 REFERENCE
 Klein, S. and Struhsberg, R.
 Direct Submission
 Submitted (02-DEC-2003) National Institutes of Health, Xenopus Gene
 Collection (XGC), National Institute of Child Health and Human
 Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
 20892-7510, USA
 NIH-MGC Project
 CONTACT: XGC help desk
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: Robert M. Grainger
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
 Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
 Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao,
 Kim Macdonald, Amara Maasson, Mike R. Mayo, Josh Moran, Ryan Morin,
 Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilia Prabhu,
 Parvaneh Seedat, JR Santos, Angelique Scherch, Ursula Skalska,
 Duane Smallus, Jeff Stoltz, Miranda Tsai, George Yang, Jacques
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAX Plate: 142 Row: m Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency OK
 analysis, similarity but not identity to protein.
 Location/Qualifiers

1..1936
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="MG75650 IMAGE:5379633"
 /csize_type="Embryo, Silurana tropicalis, tailbud (stages
 20-27)"
 /clone_1lb="NICHU_XGC_Emb7"
 /lab_host="DH10B"
 /note="vector: PCMV-SPORT6.1"
 1..1936
 /gene="MG75650"
 /db_xref="LocusID:394778"
 365..1624
 /gene="MG75650"
 /codon_start=1
 /product="hypothetical protein MG75650"
 /protein_id="AA63200.1"
 /db_xref="GI:38649365"
 /db_xref="LocusID:394778"
 /translation="MPSQSEKAPSPKSPKPVYGLVVLGNGCLPNDGRGRKRF
 LFKRKSNGVAVSSVHVSITDPAKVAISRGHISFTLSQTVARVETADKDDMF
 OIGRTSEIPDFVTVDSQNDQITQSTISPAFRIQVDRVPYARIFAAGPD
 SKNITFLGEKAKWNPDPGMDGLTNTVIVWHPGFTSESPVMEISVCGVYT
 LRRPSAORCKLVENETNIIODGSLVDCATILMPRADGLIHPKQIIFALLET
 NNAKRVQGVPLTWLGECEGFYVDADPPTPHAFPCGHVSSKSAKYSQIPLPHGTAF
 ENKRTVGGVPLTWLGECEGFYVDADPPTPHAFPCGHVSSKSAKYSQIPLPHGTAF
 HNAFCFPAIOLTGERSCVKLIFQGPVD"

ORIGIN
 Alignment Scores:
 Pred. No.: 6, 09e-164 Length: 1936
 Score: 2129.50 Matches: 384
 Percent Similarity: 96.67% Conserved: 22
 Best Local Similarity: 91.43% Mismatches: 13
 Query Match: 92.99% Gaps: 1
 DB: 5


```

Db      398 ATGTCCTCCACACACAGCGGCTCCAGGCATACGCTCCAGAGACATCACACATATCG 457
Qy      81  TTTTThleuSerArgAnGlnThrValValGluTyrThrHisApIysApThrAsp 100
      458 TACACCTTTCACGAGCAGCAGCAGGTAAGTGTGAGTATACACACACATTAAGACACAGAC 517
Db      101 MetPheGlnValGlyArgSerThrglyuserProIleAspPheValValThrAspThrIle 120
      518 ATGTTTCAGGTGGCAGGTCAACAGAAAGCCCATTTGACTCTGTGTCAACAGCAGGCTT 577
Qy      121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrglyuserThrIleSerArgPheAla 140
      578 TCCGGGGGTCAAGACCAAGAT---GCCAGATCACACAGACCACTCTAGGTTCCGA 634
Db      141 CybArgGleValCybAspArgAnGlnProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
      635 TCCAGGATCGTGTGTGACAGAAAGCAGCATACACAGCAGCATATTCGCGCAGAGATTC 694
Qy      161 AspSerSerIysAsnIlePheLeuGlyGlyValAlaAlaIleThrIysAsnProAspGly 180
      695 GATCTTCCAAAATAATCTTTCTTGAGAGAAAGCAGCAAAATGAAAAACCTGATGGA 754
Db      181 HisMetAspGlyLeuThrThraAngIyValIleuValMetHisProArgGlyGlyPheThr 200
      755 CACATGGATGACTCATCAACACGGTCTCTGTGTATCATACCCCGCAAGAGGCTTCACC 814
Qy      201 GluGlySerGlnProGlyValIleThrArgGluIleSerValCybGlyAspValIleThrLeu 220
      815 GAGGAATCCACAGCTCGAGTCTGAGAGAGATCTCTGTCTGTGGGATGTATACACCTTG 874
Db      221 ArgGlyThrArgSerAlaGlnIleArgGlyValIleuValGluSerGluThrAsnValIleu 240
      875 CAGAGAACCAAGGTCCGCCAGCAGAGGGAGAAAGCTGTGTGAAGTGAACCAACGCTCG 934
Qy      241 GlnAspGlySerLeuIleAspLeuCybGlyAlaThrLeuLeuThrArgThrAlaAspGly 260
      935 CAAGACGGTCTCTATTGACTGTGTGGGCCACTCTCTCTGAGAAACCGCAAGTGGC 994
Db      261 LeuPheHisThrProThrglnIleValIleGluAlaLeuArgGlnIleIleAsnAlaAla 280
      995 CTTTTCACGCTCTACTCAGAAAGCAGATAGAAAGCCCTCCGGCAGAGATCAATGACGC 1054
Qy      281 ArgProGlnIleCybProValGlyLeuAsnThrIleAlaPheProSerIleAsnArgIleu 300
      1055 CAAACCCAGTGGCCCGTGGGCTTAAACACCTGGGCTTCCCAAGCATCAACCGGAAGGA 1114
Db      301 ValIleGluGluIleGlnProThrAlaTyrLeuSerCybGlyHisValHisGlyTyrHis 320
      1115 GTGTGTGAAGAAAGACAGCCCTGGGCAATCTGAGCTGGGCCATGTGACCGGCTTACC 1174
Qy      321 AsnThrPglYHisArgSerAspThrgluAlaAsnGluArgGluCybProMetCybArgThr 340
      1175 CACTGGGGGCATCGGAGCAGACAGGAAGCCCAAGAGAGGAGTGTCCATGTGATGACT 1234
Db      341 ValIleProTyrValProleuThrLeuGlyCybGlyAlaGlyPheYrValAspAlaGly 360
      1235 GTGGGCGCTTACGCTCTCTGTGGCTGTGAGCGAGATTATATGCAANNNNNN 1294
Qy      361 ProPheThrHisAlaPheThrProCybGlyHisValCybSerGlyIleSerAlaIleYr 380
      1295 NNNNNNNNCATGTTTTCACCCCTGCGGGCAGCTGTCTTCAGAGAAAGCTTCCAAAGTAC 1354
Db      381 TrpSerGlnIleProleuProHisGlyThrHisAlaPheHisAlaAlaCybProPheCyb 400
      1355 TGGTGGCAGATCCCACTGCCCAAGGACGACGCGCTTCAAGCCCTGTGCGTCTGCG 1414
Qy      401 AlaThrGlnIleuValGlyGluGlnAsnCybIleValLeuIlePheGlnIleProIleAsp 420
      1415 GCCAGCGACCTGGTGTGTGAACGAACTCAATTAATTAATTTCCAAAGGTCAGAGTGC 1474

```

```

RESULT 13
AF302504
LOCUS      AF302504      1734 bp      mRNA      linear      ROD 05-JUL-2001

```

```

DEFINITION Mus musculus pellino 2 (Pell2) mRNA, complete cds.
ACCESSION AF302504
VERSION AF302504.1 GI:10242356
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE
AUTHORS Resch, K., Jockusch, H. and Schmitt-John, T.
TITLE Assignment of homologous genes, Pell1/Pell1 and Pell2/Pell2, for the Pelle adaptor protein Pellino to mouse chromosomes 11 and 14 and human chromosomes 2p13.3 and 14q21, respectively, by physical and radiation hybrid mapping
JOURNAL Cytogenet. Cell Genet. 92 (1-2), 172-174 (2001)
MEDLINE 21203570
PubMed 11306823
REFERENCE 2 (bases 1 to 1734)
AUTHORS Resch, K., Jockusch, H. and Schmitt-John, T.
TITLE Direct Submision
JOURNAL Submitted (05-SEP-2000) Developmental Biology and Molecular Pathology, University of Bielefeld, Universitaetsstrasse 25, Bielefeld, NRW 33615, Germany
FEATURES
source location/Qualifiers
1..1734
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
1..1734
/gene="Pell2"
218..1477
/gene="Pell2"
/note="similar to Drosophila melanogaster Pellino"
/codon_start=1
/product="Pellino 2"
/protein_id="AAG15392.1"
/db_xref="GI:10242357"
/translation="MESPQDEPSAPNKPVKYRELVLVLYGVALPNDGRGKSRFTLYKRYASGVKSTIHMTSPQASKALISRGHNSIYLSRSQTVVVEYTHDKDPMQVGRSTESIPDVFVDVTSQGNEDAOISITSPACRIYCDRNEPYARIPLAAGFDSSKNTFLBKKAKMKMPDGHMDGLTNVLVMPGGGFTESQPGVREISVCGVYTLKRTSADQKGLVSEFNVDGSLIDLCGTLTKRRADGIFHAPTOGHTALROETNAARQCEVGLNTTLAFPSINRKEVVEKQPMAYLSCGVHGHGHNHSDTANRECFMCMVTPGVPLMLCEAGFYVXXXXXHVFPICGVHSEKSAKXWSOIPLPHTHAFHNAACFCAQTQVGEQNCIKLIFGQVFD"
ORIGIN
Alignment Scores:
Pred. No.: 1,566-162 Length: 1734
Score: 2111.50 Matches: 388
Percent Similarity: 95.24% Conservative: 12
Best Local Similarity: 92.38% Mismatches: 19
Query Match: 92.21% Indels: 1
Gaps: 1
US-10-041-030-4 (1-420) x AF302504 (1-1734)
Qy      1 MetPheSerProGlyGluGlnGluHisCybAlaProAsnIleGluProValIleYrGly 20
Db      218 ATGTTTTCCTCCGCGGCAAGAGAACCCAGCCGCCCAACAGAGACCGGTGAATAACAG 277
Qy      21 GluLeuValIleuGlyTyrAnGlyValIleuProAsnIleAspArgGlyArgArgIly 40
Db      278 GAGCTGTGTGCTCTGGGGTAACAATGTCTTAACTTAATGTGTGACAGGGGCAAGAA 337
Qy      41 SerArgPheAlaLeuTyrIleArgProIyAlaAsnGlyValIleYrProSerThrValHis 60
Db      338 AGCAATTTTACCTTAAAGCGGACCTAAGCAGATGGTGTCAAAACCAAGCAATCCAC 397
Qy      61 ValIleSerThrProGlnAlaSerIyAlaIleSerCybIyGlyGlnIleSerIleSer 80
Db      398 ATGTCCTCCACACACAGCGGCTCCAGGCATACGCTCCAGAGACATCACACATATCG 457

```

```

Oy      81  TyrThrLeuSerArgAsnGlnThrValValValGluTyrThrHisAspLysAspThrAsp 100
Db      458  TACAGGTTGTACGAGCGCAGCGGTAGGTGGGTGACACACAGCAATTAACACAGC 517
Oy      101  MetPheGlnValGlyArgSerThrGluSerProIleAspPheValThrAspThrIle 120
Db      518  ATGTTTCAGGTGGGAGGTCAACAGAAAGCCCATTTGACTTGCTGTGTCTCACACACGGTT 577
Oy      121  SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      578  TCCCGCGGTGACAGAAAGAAAGAT---GCCCGAGATCACACAGAGACCATCTCTGTGTCGA 634
Oy      141  CysArgGlnIleValCysAspArgAsnGlnLysProTyrThrTrpAlaArgIlePheAlaIleGlyPhe 160
Db      635  TGCAGAGATGTGTGTGACAGGAAACAGCATACACAGCAGCATATTGCGGCGAGATTC 694
Oy      161  AspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTrpLysAsnProAspGly 180
Db      695  GATTCTTCCAAATAATATCTTTCTTGTGAGAGAAAGCAGCAAAATGGAATAACCTGATGA 754
Oy      181  HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db      755  CACATGGATGGACTCACTACCAACGAGTGTCTGTGTGATGACCCGCAAGAGAGCTTCA 814
Oy      201  GluGluSerGlnProGlyValTrpArgGluIleSerValCysGlyLysAspValTyrThrLeu 220
Db      815  GAGGAAATCCAGCCCTGGAGTCTGGAGAGAGATCTCTGTGTGGGAGTGTATACCTTGC 874
Oy      221  ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGlyLysThrAsnValLeu 240
Db      875  CGAGAGACCAAGTCCGGCCAGCAGAGGGGAAAGCTGTGTGAAGAGAGACCAAGCTCTGC 934
Oy      241  GlnAspGlySerLeuIleAspLeuGlyValAlaThrLeuLeuTrpArgThrAlaAspGly 260
Db      935  CAAGCGGCTCCCTCATTTGACCTGTGTGGGCGACTCTCTCTGTGAGAAACCCCAATGGCC 994
Oy      261  LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaIle 280
Db      995  CTTTTCACGCTCTCACTCAAGAGACATAGAGAGCCCTCCGCGAGAGATCAATGACGCC 1054
Oy      281  ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
Db      1055  CGACCCCGAGTGGCCCGTGGGCTTAAACACCTGGGCTTCCCGACAGATCAACCGAGAGGA 1114
Oy      301  ValValGluGluLysGlnProThrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
Db      1115  GTGGTGAAGAGAAACAGCCCTGGGCAATCACTGAGCTGGGCGCATGTGCACGCTAACAC 1174
Oy      321  AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
Db      1175  CACTGGGGCGCATCGAGCGACAGCAGAGCAAGCAAGAGGAGATGCCATGTGCTAGTACT 1234
Oy      341  ValGlyProTyrValProLeuTrpLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
Db      1235  GTGGGCGCCCTAGCTCTCTCTGTGGGCTGTGAGCGAGATTTTATGTGANNNNNN 1294
Oy      361  ProProThrHisAlaPheThrProCysGlyHisValCysSerGlnLysSerAlaLysTyr 380
Db      1295  NNNNNNNCTCATGTTTACCCCTGGCGGCGAGCTGTTCGAGAAAGCTGCGCAAGTAC 1354
Oy      381  TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db      1355  TGTGTGCGAGATCCACTGCGCCCAAGAGAGCAGCGGTTTTCAGCGCGCTCTCTGTC 1414
Oy      401  AlaThrGlnLeuValGlyGluGlnAsnGlyIleLysLeuIlePheGlnGlyProIleAsp 420
Db      1415  GCCACGACAGTGTGTGTGGAACAGAACTGCATCAATTGATTTCCAAGGTCCAGTGGAC 1474

```

```

RESULT 14
LOCUS      BC075973          2474 bp      mRNA      linear      VRT 27-JUL-2004
DEFINITION Danio rerio zgc:92268, mRNA (cDNA clone MGC:92268 IMAGE:7053197),

```

```

complete cds.
BC075973
BC075973.1 GI:49903830
MGC.
Danio rerio (zebrafish)
Danio rerio
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 2474)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,W.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uebli,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahney,J., Hellton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Kizylinski,M.I., Skalska,U., Smalins,D.E.,
Scherer,A., Schein,J.E., Jones,S.J., and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
1247932
2 (bases 1 to 2474)
Strausberg,R.
Direct Submission
Submitted (06-JUL-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IPAK Plate: 179 Row: m Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1..2474
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:92268 IMAGE:7053197"
/tissue_type="whole, adult zebrafish"
/clone_lib="NIH_ZGC_7"
/lab_host="DH10B"
/note="vector: pXpress1"
1..2474
/gene="zgc:92268"
/db_xref="locusID:436899"
218..1504
/gene="zgc:92268"

```

```

/codon_start=1
/product="zgc:92268"
/protein_id="AAH75973.1"
/db_xref="GI:49903631"
/translation="MFSPOEHCAPTKEPVKKGLVVLGVNGLPNGRGRKSRFA
LYRBAKNGVKSSTVHLIANTPOASKVANKGQHSISYTSRNTQVYVEVSHDPTMF
QGRSTREPDIPIVVDITVSGGSEETPTOSTISFPACKVCEKRPYTAIRPAAGF
DSSKNIPLUEKAKMKNPDPGHMDGLTTNGVLVHMPRGFTESKPMVEISVCGDVY
TLRETSAGTRGLVSESNVLQDGLVLCATLMTADGLFHTPTQKHLALOE
LNAARPOCPLVNTLAFPSMORCSALSPADQDQPVYLAIGHVGHMDGHSB
RESNORCEPCMRVAVPYVPLMGCEPAFYVDGAPTHFVPCGHSCEKSRVYAEI
PLPHGTHAHACPCFATQTLNLTQKAKLIPOGPID"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 9,59e-154 Length: 2474
Score: 2006.00 Matches: 361
Percent Similarity: 90.89% Conservative: 28
Best Local Similarity: 84.35% Mismatches: 31
Query Match: 87.60% Indels: 8
DB: 5 Gaps: 2

```

US-10-041-030-4 (1-420) x BC075973 (1-2474)

```

QY 1 MetPheSerProGlyGlnGluGluHsCybAlaProAaenlyegluProVallyrGly 20
DB 218 ATGTTTTCACCGGGTCAAGAGAGCACTGTGCGCCCAACCAAGAGCAGTAAAGTACGGG 277
QY 21 GluLeuValValLeuGlyTyraenGlyAlaLeuProAaenlyAaPrgGlyAaGly 40
DB 278 GAACCTGGTGTTCATGATTCATGCTCCTCCCAATGATGACAGAGGCGCAAGAA 337
QY 41 SerArgPheAlaLeuTyrylAaPrgProlyAaAaenGlyVallyrProSerThraValHs 60
DB 338 AGCCGGTTTGCCCTTTACAAGAGAGCCAAAGCCAAATGGGTCAAGCCAGCACTGACAC 397
QY 61 ValIleSerThraProGlnAlaSerlyAaIleSerCyalyeglyGlnHsSerIleSer 80
DB 398 ATCTTACACACACACACACACACACACACACACACACACACACACACACACACAC 457
QY 81 TyrThraSerArgaenGlnThraValValGluTyThraHsAaPrgAaPrgThraP 100
DB 458 TACACGCTTCCAGAAACCAAGAGGTGTGTGAGTACACCATACAAAGACACAGAC 517
QY 101 MetPheGlnValGlyAaSerThraGlySerProIleAaPrgPheValValThraPThra 120
DB 518 ATGTTCCAGATCGCGCATCGACGAGGGCGCATTTGCTGTGATGACACTGTC 577
QY 121 SerGlySerGlnAaenThraPrgGluAlaGlnIleThraGlnSerThraIleSerArgPheAla 140
DB 578 TCAAGGGGAGAGAGGTGAGAGACACCATCATCTCAAGACACATCTCCCGCTTCGCG 637
QY 141 CybArgIleValCybAaPrgAaenGluProTyThraAlaGlyIlePheAlaIleGlyPhe 160
DB 638 TCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 697
QY 161 AaPrgSerlyAaenIlePheLeuGlyGluTyrylAaAlaIleTyThraPrgAaPrgGly 180
DB 698 GATTTCTCAAGAAATATATTTCTCGAAGAAAGACCCAAATGAGAAATCCAGATGAT 757
QY 181 HsAaPrgGlyLeuThraHsAaenGlyValLeuValMetHsAaPrgGlyGlyPheThra 200
DB 758 CACATGTGACGTTTGTCTCAATGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 817
QY 201 GluGluSerGlnProGlyValTyraGlyIleSerValCybGlyAaPrgValTyThraLeu 220
DB 818 GAGGAATCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 877
QY 221 ArgGluThraSerGlnAlaGlnIleAaPrgGlyTyrylAaLeuValGlySerGluThraAaVal 240
DB 878 AGAGAAACACGCTGTCTCAACTCGAGGAAATTTGTGTGAGAGTGAAGTAACTGTCTG 937

```

```

QY 241 GlnAaPrglySerIleuAaenPleuCybGlyAlaAthrLeuLeuThraPrgThraAaPrgly 260
DB 938 CAGAGAGGCTCTCTGTGTACCTTTGTGAGAACCCCTCTGTGGGACACCGCTGACGCG 997
QY 261 LeuPheHsThraProThraGlnlyHsIleGluAlaLeuAaPrgGlnGluIleAaenAla 280
DB 998 CTTTTCACACCCCCACACAGAACACCTGAGGCTTTCGCGAGAGCTGAACGACGCC 1057
QY 281 ArgProGlnCybProValGlyLeuAaenThraAaPheProSerIleAaenArg----- 298
DB 1058 CGGCCCCAGTGCCTGTGTGGCTTCAACACTGCGCTTCCAGCATGACGAGCTGCAGT 1117
QY 299 -----LysGluValValGluGluTyrylAaenProThraAlaTyThraLeuSerCy 313
DB 1118 CGTGCCCTGAGACGCTCCCTGCCAGAAACAAACAGCCCTGGGTCTACCTTGCCTGT 1177
QY 314 GlyHsValHsGlyTyThraHsAaenThraPrglyHsAaPrgAaPrgThraGlnAaen---Glu 332
DB 1178 GGTATGTTTCATGTGCTACACACAGCTGGGACACCATGACAGAGGAGTCCAAATGCCCA 1237
QY 333 ArgGluCybProMetCybAaPrgThraValGlyProTyrylAaProLeuThraPrgGlyCybGlu 352
DB 1238 AGAGAGTGCCTCATGTGACAGGCGCGGTCCCTATGTGCACTGTGCTGCGATGTGA 1297
QY 353 AlaGlyPheTyrylAaPrgAlaGlyProProThraHsAaPheThraProCybGlyHsVal 372
DB 1298 CCAGCTTTCATGATGACACCTGTGTGCGCCACACACAGTCTTGTGATCTGTGTGCGCAT 1357
QY 373 CybSerGluTyrylAaPrgThraPrgSerGlnIleProLeuProHsIleGlyThraHsAla 392
DB 1358 TGTCTGAAAGATCCACCGGCTATGTGGCAGAGATTCCTTTGCCATGTGACACCGCT 1417
QY 393 PheHsAlaAlaCybProPheCybAaPrgThraGlnLeuValGlyGluAaenCybIleTy 412
DB 1418 TTCACGCTGCTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1477
QY 413 LeuIlePheGlnGlyProIleAaPrg 420
DB 1478 CTCATCTTTCAGGGGCTATGAC 1501

```

RESULT 15

AR483566

LOCUS AR483566 1257 bp DNA linear PAT 14-MAY-2004

DEFINITION Sequence 3 from patent US 6703487.

ACCESSION AR483566

VERSION AR483566.1 GI:47246650

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE

1 (bases 1 to 1257)

AUTHORS Bird,T.A. and Cosman,D.J.

TITLE Humanellino polypeptides

JOURNAL Patent: US 6703487-A 3 09-MAR-2004;

FEATURES

location/Qualifiers

1..1257

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 7.63e-147 Length: 1257

Score: 1917.00 Matches: 342

Percent Similarity: 90.48% Conservative: 38

Best Local Similarity: 81.43% Mismatches: 38

Query Match: 83.71% Indels: 2

DB: 6 Gaps: 1

US-10-041-030-4 (1-420) x AR483566 (1-1257)

QY 1 MetPheSerProGlyGlnGluGluHsCybAlaProAaenlyegluProVallyrGly 20

DB 1 ATGTTTTCCTGATCAAGAAATATCAT-----CATCTAAAGACACAGTAAATATGTGT 54

```

QY 21 GILLeuValIValLeuGIYThrAsnGIYAlaLeuProAsnGIYAspArgIYArgArgIYs 40
Db 55 GAACTCATGTGTTGGATATATGATGATCTCCCAAGCGTGATATGAGAAAGAGGAAA 114
QY 41 SerArgPheAlaLeuTYrLYsArgProLYsAlaAsnGIYValLYsProSerThrValHis 60
Db 115 AGTAGGTTTGGCTTGTGTTTAAAGACTTAAGCAATGGGGTGAAGCCAGCACTGGCAT 174
QY 61 ValILeSerThrProGlnAlaSerLYsAlaILeSerCYsLYsGIYGlnHisSerILeSer 80
Db 175 ATTGCTTGACTCTCAGGCTGCAAGCAAGCAATAGCAACAAAGACCAAGCATGACATATCA 234
QY 81 TYrThrLeuSerArgAsnGlnThrValValIGluTYrThrHisAspLYsAspThrAsp 100
Db 235 TATACTTATCTCGGGCCCAAGACTGTGGTGTGTAATATCTCATGACAGCAACCCGAT 294
QY 101 MetPheGlnValIGLYsSerThrGluSerProILEAspPheValValThrAspThrILE 120
Db 295 ATGTTTCAGATTGGCCGGTCGACTGAAAGCCCATGATTTTGTAGTACTGACAGGTT 354
QY 121 SerGIYSerGlnAsnThrAspGluAlaGlnILEThrGlnSerThrILESerArgPheAla 140
Db 355 CCTGGAAGCAAGTAATTCTGATACACAGTCAGTACAAAGCATATATCAAGATTTCGCC 414
QY 141 CYsArgILEValCYsAspArgAsnGluProTYrThrAlaArgILEPheAlaIAGLYPhe 160
Db 415 TGCAAAATCATATGTGAAGGAAATCTCCCTTACAGCAAGGATTTATGCTGACAGGTTT 474
QY 161 AspSerSerLYsAsnILEPheLeuGIYGLYsValAlaLYsTrpLYsAsnProAspGIY 180
Db 475 GACTCATCAAAAACATCTTCTTGGGGGAAGGCTGCCAAATGGAAAGACATCAAGATGGA 534
QY 181 HisMetAspGIYLeuThrThrAsnGIYValILEValMetHisProArgGIYGLYPheThr 200
Db 535 CAGATGGATGGCTTAGCACTAATGGTCTTGATGATCATCCAGCAATGGGTTTCACA 594
QY 201 GILGluSerGlnProGluValITrpArgGluILESerValCYsGIYAspValTYrThrLeu 220
Db 595 GAAAGACTCCAAAGCTCGAATATGAGAAAGAAATATCGGTGTGGAAATGATTTAGCTTA 654
QY 221 ArgGluThrArgSerAlaGlnGlnArgGIYLYsLeuValIGluSerGluThrAsnValLeu 240
Db 655 CGTGAAACCAAGATCGGCTCAGCAGAGAGAAATGTGTGAATTTGAACCAATCAGTTA 714
QY 241 GlnAspArgLYSerLeuILEAspLeuCYsGIYAlaITHrLeuLeuITrpArgThrAlaAspGIY 260
Db 715 CAAGATGGCTGTTAATTGACCTGTGTGGCAACATTTGTATGGGTACTGCAGAAAGGC 774
QY 261 LeuPheNIEThrProThrGlnLYsHisILEGluAlaLeuArgGlnGluILEAsnAlaAla 280
Db 775 CTTTCCCAACACTCTACCCGTGAAGCATTTAGAAGCTTTAAGACAGAAATCAATGCAGCA 834
QY 281 ArgProGlnCYsProValIGLYsLeuAsnThrLeuAlaPheProSerILEAsnArgLYsGlu 300
Db 835 CCACTCAGTGGCCCTGTAGGTTTCAACACATAGCATTTCTAGTATGAAGAGAAAGAC 894
QY 301 ValILEGluGlnLYsGlnProITrpAlaTYrLeuSerCYsGIYHisValHisGLYTYrHis 320
Db 895 GTTGTAGATGAAGAAACAAACATAGGTATATCTAAACTGTGGCCATGTACATGGCTATCAT 954
QY 321 AsnITrPGLYHisArgSerAspThrGluAlaAsnGluArgIUCysProMetCYsArgThr 340
Db 955 AACTGGGGAACAAGAAAGACAGTAGTAGCCAAAGATCGGAATGCTCTATGTAGGTCT 1014
QY 341 ValIGLYProTYrValProLeuITrPLeuGIYCYsGluAlaGLYPheTYrValAspAlaGLY 360
Db 1015 GTTGTGCTCCATGTCTGTGTGGCTTGAGATGAAAGTAGATTTATGTGACCGCGGC 1074
QY 361 ProProThrHisAlaPheThrProCYsGIYHisValCYsSerGluLYsSerAlaLYsTYr 380
Db 1075 CCTCAACCCATGCGTTAGCCGCTGTGGCAATGTGTTCAGAAAAGCAACTGCTAT 1134

```

```

QY 381 TrpSerGlnILEProLeuProHisGIYThrHisAlaPheHisAlaAlaCYsAspProPheCYs 400
Db 1135 TGATCCAGATCCCACTTCTCATGTACTCATCTTTTCAATGACGCTGTCTTTGT 1194
QY 401 AlaThrGlnLeuValIGLYGlnGlnAsnCYsILELYsLeuILEPheGlnGIYProILEAsp 420
Db 1195 GCACATCAGTTGGCTGTGGAACAAAGCTTACATCAGACTTATTTTCAAGAACCTCTAGAC 1254

```

Search completed: December 9, 2004, 15:53:39
 Job time : 7563 secs


```

Db      61 GAGCTGTGTGTGGTGAACAATGGTCTTACCAATGGAGATAGAGAGCGAGAAA 120
Qy      41 SerArgPheAlaLeuTyrIlysaGrPolyAlaAsnGlyValIysProSerThrValHis 60
Db      121 AGTAGATTTGGCCCTTACCAAGCGGCCCAAGGCGAAATGGTGTCMAACCCGACGCCCTCAT 180
Qy      61 ValIleSerThrProGlnAlaSerIlyValIleSerCysIysGlyGlnHisSerIleSer 80
Db      181 GTGATATTCACGCCCCAGGCGATCAAGGCTATACGTTGCAAGGTCAACACATATATCC 240
Qy      81 TyrThrLeuSerArgAsnGlnThrValValIleGluTyrThrHisAspIysAspThrAsp 100
Db      241 TACACTTTGTCAAGAAATCAGACTGTGTGTGTAATACACATATAGAGATAGAT 300
Qy      101 MetPheGlnValIleYarSerThrGluSerProIleAspPheValIleThrAspThrIle 120
Db      301 ATGTTCAGGTGGCGAGATCAACAGAAAGCCCTATCGACTTCGTTGTCAACAGACAGATT 360
Qy      121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      361 TCTGGACGCCAGAACCGAGACAGAGCCAGATCACAGACACATATCCAGGTTCCGC 420
Qy      141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
Db      421 TCGAGATCGTGTGGCAGAGAAATGAACCTTACAGACAGGATATTCGCCCGCGATT 480
Qy      161 AspSerSerIlyAsnIlePheLeuGlyGlyValAlaAlaIlystTplyAsnProAspGly 180
Db      481 GACTCTTCCAAAACATATTTCTTGGATGAAGCGCAAGCGCAAAAGCGGAAACCCCGACGCG 540
Qy      181 HisMetAspGlyLeuThrThrAsnGlyValIleValMetHisProArgIlyGlyPheThr 200
Db      541 CACATGGATGGGCTCACTAATGAGCGTCTGTGATGATGATTCACAGAGGGGCTTCAC 600
Qy      201 GlnGluSerGlnProGlyValIleTparGlnIleSerValCysGlyAspValTyrThrLeu 220
Db      601 GAGGATCCCAAGCCCGGGGTGGCGGAGATCTGTCTGTGAGATGTATACCTTG 660
Qy      221 ArgGluThrArgSerAlaGlnGlnArgIlyIlyLeuValIleGluSerGluThrAsnValLeu 240
Db      661 CGAGAAACCGAGTGGCGCCAGCAACAGAGAAAGCTGTGTGAAGTGAAGCAACGCTCG 720
Qy      241 GlnAspGlySerIleLeuAspLeuCysGlyValIleThrLeuThrArgThrIleAspGly 260
Db      721 CAGGACGCTCCCTCTTAACCTGTGTGGGCCACTCTCTCTGGAGAACAGAGTGG 780
Qy      261 LeuPheHisThrProThrGlnIlyHisIleGluAlaLeuArgGlnIlyIleAsnAlaIle 280
Db      781 CTTTTTCATATCTCAACTCAGAAAGCATAGAAGCCCTCGGAGAGATTTAACCGCGCG 840
Qy      281 ArgProGlnIlyProValIleGluAsnThrIleAlaPheProSerIleAsnArgIlyGlu 300
Db      841 CGGCTTCAGTGTCTGTGTGGGCTCAACACCTGGCCCTCCAGCTCAACAGAGAAAG 900
Qy      301 ValIleGlnIlyGluSerIlyProTparIleTyrLeuSerCysGlyHisIleGlyTyrHis 320
Db      901 GTGGTGAAGAGAAAGAGCCCTGGGCAATTCAGTTGTGGCCACCTGACCGGATCCAC 960
Qy      321 AsnTrrGlyHisIleArgSerAspThrGluAlaAsnGluArgIlyCysProMetCysArgThr 340
Db      961 AACTGGGGCCATCGAGGTACACGAGGCGCAACAGAGAGGAGTGTCCATGTGACGAGAT 1020
Qy      341 ValIlyProTyrValIleProLeuThrLeuGlyCysGlyAlaGlyPheTyrValIleAspAlaGly 360
Db      1021 GTGGGCCCCCTAATGTGCTCTGTGGCTGTGAGCGAGATTTATGTACACGACGAG 1080
Qy      361 ProProThrHisAlaPheThrProCysGlyHisIleValCysSerGlnIlySerAlaIlystTrr 380
Db      1081 CGGCCACATATGCTTCACTCCCTGTGACACGATGTGCTCGAGAGAGTCTGDAADATAC 1140
Qy      381 TrrSerGlnIleProLeuProHisIlyThrHisAlaPheHisAlaIleCysProPheCys 400

```

```

Db      1141 TGTCTCAGATCCGTTGCTCATGGAATCATGATTTACGCTGTCCCTTCTGT 1200
Qy      401 AlaThrGlnLeuValIleGlyGlnIlyAsnGlyIleIlyLeuIlePheGlnGlyProIleAsp 420
Db      1201 GCTACAGACGTGTGGGAGAGCAAAATGCATCAAAATTAATTTTCAAGGTCAATTGAC 1260

RESULT 2
US-09-843-905A-5
/ Sequence 5, Application US/09843905A
/ Patent No. 6703487
/ GENERAL INFORMATION:
/ APPLICANT: Bird, Timothy A.
/ APPLICANT: Cosman, David J.
/ TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
/ FILE REFERENCE: 2990-A
/ CURRENT APPLICATION NUMBER: US/09/843,905A
/ PRIOR FILING DATE: 2001-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,198
/ PRIOR FILING DATE: 2000-04-28
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 5
/ LENGTH: 1260
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-843-905A-5

Alignment Scores:
Pred. No.: 2,066-231 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.38% Conservative: 10
Best Local Similarity: 95.00% Mismatches: 10
Query Match: 95.00% Indels: 1
DB: Gaps: 1

US-10-041-030-4 (1-420) x US-09-843-905A-5 (1-1260)
Qy      1 MetPheSerProGlnGlnGlnIlyHisCysAlaProAsnIlyGluProValIlyTrrGly 20
Db      1 ATGTTTCCCGGGCCAGAGAGAACCCGCGCCCAACAGAGACCGGTAATATACGG 60
Qy      21 GlnLeuValIleGluTyrAsnGlyValIleuProAsnGlyAspArgGlyYarArgIlyAs 40
Db      61 GAGCTGTGTGTCTGGGTACATATGTTCTTACTTAATGGTGAAGAGGGGCGAGGAAA 120
Qy      41 SerArgPheAlaLeuTyrIlysaGrPolyAlaAsnGlyValIysProSerThrValHis 60
Db      121 AGCAGATTTGGCCCTTATAGAGCGACCTACGCTGATGTCMAACCCAGCACATATCAC 180
Qy      61 ValIleSerThrProGlnAlaSerIlyValIleSerCysIysGlyGlnHisSerIleSer 80
Db      181 ATGGTCTCACACACAGCGCGTCCAGGCGCATACGCTCCAGAGGACATCACACATATCG 240
Qy      81 TyrThrLeuSerArgAsnGlnThrValValIleGluTyrThrHisAspIlyAspThrAsp 100
Db      241 TACACGTTGTACAGGAGCGACAGCGGTATGAGTGAATACACACAGATTAAGACACGAGAC 300
Qy      101 MetPheGlnValIleYarSerThrGluSerProIleAspPheValIleThrAspThrIle 120
Db      301 ATGTTCAGGTGGCGAGTCAACAGAAAGCCCTATCGACTTCGTTGTCAACAGACAGATT 360
Qy      121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      361 TCCGCGGCTCGAAGAAAGAT--GCCCAAGATCACAGACACACCATCTTAGGTTCCGA 417
Qy      141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
Db      418 TCGACAGATCGTGTGACAGAGAACAGCAATACAGACCCATATTCGCGGAGGATTC 477
Qy      161 AspSerSerIlyAsnIlePheLeuGlyGlyValAlaAlaIlystTplyAsnProAspGly 180
Db      478 GATTCCTCAAAAATATCTTCTTGTGAGAGAAACAGCAAAATGGAAGAAACCCCTGATGA 537

```

QY 181 HlmeAaBpGlyLeuThrThraBnglyValLeuValMehiBProAaBpGlyGlyPheThr 200
 DB 538 CACATGATGAGTCACTACCAATGATGCTTATGATGACCCGAGAGAGCTTCACC 597
 QY 201 GluGluSerGlnProGlyValTPAaGluIleSeValCyBglYaaPValTYrThrLeu 220
 DB 598 GAGAAATCCAGCTGAGTCTGAGAGAGATCTGTCTGTGGGATGTGACCTTG 657
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLeuValGluSerGluThrAenValLeu 240
 DB 658 CGAGAAACCAAGTCCGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717
 QY 241 GlnAaBpGlySerLeuIleAaBpLeuCyBglYalThrLeuLeuTPaArgThrAlaAaBpGly 260
 DB 718 CAGAGAGGCTCCCTCATTCAGCTGTGTGGGCACTCTCTCGAGAGAGAGAGAGAG 777
 QY 261 LeuPheHlThrProThrGlnYbHlAlaGluAlaLeuAaBpGlnGluIleAaBpAla 280
 DB 778 CTTTTCAGCTCTCTCATTCAGAGACATGAGAGAGAGAGAGAGAGAGAGAGAGAG 837
 QY 281 ArgProGlnCyBpProValGlyLeuAaBpThrLeuAlaPheProSerIleAaBpGlyGlu 300
 DB 838 CGAGCCCAAGTCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897
 QY 301 ValValGluGluLeuGlnProTPaIaTYrLeuSerCyBglYalHlAlaHlGlyTYrHlAla 320
 DB 898 GTGGTGAAGAAG 957
 QY 321 AaBpThrGlnHlAaBpSerAaBpThrGluAlaAaBpGluCyBpProMetCyBpArgThr 340
 DB 958 AGCTGGAGGAG 1017
 QY 341 ValGlyProTYrValProLeuTPaLeuGlyCyBglYalAaBpGlyPheTYrValAaBpAlaGly 360
 DB 1018 GTGGGAG 1077
 QY 361 ProProThrHlAlaPheThrProCyBglYalHlAlaCyBpGluYbSerAlaTYr 380
 DB 1078 CCCCCAATCAGCTTCATCCCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137
 QY 381 TrpSerGlnIleProLeuProHlAlaTYrHlAlaPheHlAlaAlaCyBpProPheCyB 400
 DB 1138 TGGTCCAGATCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
 QY 401 AlaThrGlnLeuValGlyGluGlnAaBpGlyIleYbSerLeuIlePheGlnGlyProIleAaBp 420
 DB 1198 GCCAGCAGCTGTGTGTGAACGAAGTCAATGATTTTCCAGAGTCCAGAGTGC 1257

RESULT 3
 US-09-843-905A-3
 / Sequence 3, Application US/09843905A
 / Patent No. 6703487
 / GENERAL INFORMATION:
 / APPLICANT: Bird, Timothy A.
 / APPLICANT: Cosman, David J.
 / TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 / FILE REFERENCE: 2990-A
 / CURRENT APPLICATION NUMBER: US/09/843,905A
 / PRIOR FILING DATE: 2001-04-27
 / PRIOR APPLICATION NUMBER: US 60/200,198
 / NUMBER OF SEQ ID NOS: 15
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO 3
 / LENGTH: 1257
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / US-09-843-905A-3

Alignment Scores:
 Pred. No.: 9,896-203 Length: 1257
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.48% Conservative: 38

Best Local Similarity: 81.43% Mismatches: 38
 Query Match: 83.71% Indels: 2
 DB: 4 Gaps: 1

US-10-041-030-4 (1-420) x US-09-843-905A-3 (1-1257)

QY 1 MetPheSerProGlyGlnGluGlnHlAlaPheProAaBpGluProValIleTYrGly 20
 DB 1 ATGTTTCCCGATCAAGAAAAATCAT-----CAATCAAGAGAGAGAGAGAGAGAGAG 54
 QY 21 GluLeuValIleGluTYrAaBpGlyAlaLeuProAaBpGlyAaBpArgGlyAaBpArgGly 40
 DB 55 GAACTCATTCCTTAGGATTAATGATCTCTCCAAAGAGAGAGAGAGAGAGAGAGAGAG 114
 QY 41 SerArgPheAlaLeuTYrIleAaBpProIleAaBpGlyValIleAaBpSerThrValHlAla 60
 DB 115 AGTAGTTCCTTTGTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 174
 QY 61 ValIleSerThrProGlnAlaSerIleValIleSerCyBglYalGlnHlAlaSerIleSer 80
 DB 175 ATGCTTGTATCTCAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 234
 QY 81 TYrThrLeuSerArgAaBpGlnThrValValIleGluTYrThrHlAaBpIleAaBpThrAaBp 100
 DB 235 TATACCTTATCTCGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 294
 QY 101 MetPheGlnValGlyAaBpSerThrGluSerProIleAaBpPheValIleThrAaBpThrIle 120
 DB 295 AGTTTCAAGATTCAG 354
 QY 121 SerGlySerGlnAaBpThrAaBpGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 355 CTTGAG 414
 QY 141 CyBpGlnIleValCyBpAaBpAaBpGlnProTYrThrAlaAaBpIlePheAlaAlaGlyPhe 160
 DB 415 TGCAGAAATATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474
 QY 161 AaBpSerSerIleAaBpIlePheLeuGlyGluYbAlaIleTYrIleAaBpAaBpAaBpGly 180
 DB 475 GACTATCAAAAAACATCTTTCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 534
 QY 181 HlmeAaBpGlyLeuThrThraBnglyValLeuValMehiBProAaBpGlyGlyPheThr 200
 DB 535 CAGATGAG 594
 QY 201 GluGluSerGlnProGlyValTPAaGluIleSeValCyBglYaaPValTYrThrLeu 220
 DB 595 GAAAGCTCCAGAGCTGAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLeuValGluSerGluThrAenValLeu 240
 DB 655 CGTGAAG 714
 QY 241 GlnAaBpGlySerLeuIleAaBpLeuCyBglYalThrLeuLeuTPaArgThrAlaAaBpGly 260
 DB 715 CAGAGAGGCTGTGTATTAATCACTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
 QY 261 LeuPheHlThrProThrGlnYbHlAlaGluAlaLeuAaBpGlnGluIleAaBpAla 280
 DB 775 CTTTCCACATCTCTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 834
 QY 281 ArgProGlnCyBpProValGlyLeuAaBpThrLeuAlaPheProSerIleAaBpArgGlyGlu 300
 DB 835 CGAGCTCAGTGCCTGTAGGTTCAACACATGACATTTCTTAATGAGAGAGAGAGAGAG 894
 QY 301 ValValGluGluLeuGlnProTPaIaTYrLeuSerCyBglYalHlAlaHlGlyTYrHlAla 320
 DB 895 GTTGTAG 954
 QY 321 AaBpThrGlnHlAaBpSerAaBpThrGluAlaAaBpGluAaBpGlyCyBpProMetCyBpArgThr 340
 DB 955 AACTGGAG 1014

```
QY 341 ValGlyProTyrValProLeuTribLeuGlyCYSGluAlaGlyPheTyrValAspAlaGly 360
| | | | |
Db 1015 GTTGGCCCTCATGTTCTCTGTGGCTGGATGTGAAGCTGGATTTTATGTGAGCCGCC 1074
QY 361 ProProThriAlaAlaPheThrProCYSGlyHisValCYsSerGluLysSerAlaLysTyr 380
| | | | |
Db 1075 CTTCCCAACCATGCGTTTGAAGCCGTGTGGCATGTGTGTTCAGAAAAAGCAACTGCCAT 1134
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCYsProPheCYs 400
| | | | |
Db 1135 TGGTCCCAATGCCACTTCTCTCATGTGACTCATATTTTCATGACCTGTCCCTTTGT 1194
QY 401 AlaThrGlnLeuValGlyGluGlnAncySileLysLeuIlePheGlnGlyProIleAsp 420
| | | | |
Db 1195 GCACATCACTGGCTGTGTGAACAAGCGTACATCAAGCTTATTTTCAAGAGACTTGAAC 1254

RESULT 4
US-09-843-905A-1
; Sequence 1, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-843-905A-1

Alignment Scores:
Pred. No.: 3,55e-202 Length: 1257
Score: 1912.00 Matches: 341
Percent Similarity: 90.48% Conservative: 39
Best Local Similarity: 81.19% Mismatches: 38
Query Match: 83.49% Indels: 2
DB: 4 Gaps: 1

US-10-041-030-4 (1-420) x US-09-843-905A-1 (1-1257)
QY 1 MetPheSerProGlyGlnGluHisCyAlaProAsnLysGluProValLysTyrGly 20
| | | | |
Db 1 ATGTTTCTCTCATGATCAAGAAATATCAT-----CCTTCCAAAGCCCAAGTAAATATGCG 54
QY 21 GluLeuValValLeuGlyLysThrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGly 40
| | | | |
Db 55 GAACCATATGTCTTAGATATATATGATCTCTCCCAACGCGTATGAGAGAGAGAGAAA 114
QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
| | | | |
Db 115 AGTAGTTTCTTTGTTTAAAGACCTTAAGCAATGGGCTGAAGCTTACCGCTGAC 174
QY 61 ValLysSerThrProGlnAlaSerLysAlaIleSerCyAlaGlyGlnHisSerIleSer 80
| | | | |
Db 175 ATTGATGATGATCTCTCAAGCTGCCAAGCAATAGCAACAAGACCATATGATATCA 234
QY 81 TyrThrLysSerArgAsnGlnThrValValValGlyTyrThrHisAspLysAspThrAsp 100
| | | | |
Db 235 TATACTTATATCTGACCCCAAGACGGTGTGATATATATCATGACACGAACTGAT 294
QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValValThrAspThrIle 120
| | | | |
Db 295 ATGTTTCAGATTGTGTGCGTCAACTGAAGTCTTATGATTTTGTAGTAACTGACACCGTT 354
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
```

```
Db 355 CTTGGAAGTCAGAGTAATTCGACACGACGTCAAGATACAAAGCATATATCAAGATTGCC 414
| | | | |
QY 141 CybArgGlyValCYsAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaLysPhe 160
| | | | |
Db 415 TGTAGATCAATATGTGAAGCGAGTCCCTTTTACGCTCGGATTTATGCTGACAGGTTT 474
QY 161 AspSerSerLysAsnIlePheLeuGlyGluLysValAlaLysTrpLysAsnProAspGly 180
| | | | |
Db 475 GATTATCAAAAAACATCTTTCTTGGGAGAAAGCTGCCAAGTGAAGACATCTGATGG 534
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
| | | | |
Db 535 CAGATGATGCTGCTTACCACTTAATGAGTCTTGTGATGATCATCCAGTAATGGGTTACA 594
QY 201 GluGlnSerGlnProGlyValITPArgGluIleSerValCYsGlyAspValTyrThrLeu 220
| | | | |
Db 595 GAAGACTCCAAACCTCGAATATGAGAGAAATATCATGTGTGGAAATGTCTTCAGTGTG 654
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGlnSerGluThrAsnValLeu 240
| | | | |
Db 655 CGTGAACCAATCAATCGCTCAGCAGAGAGAAAGATGTGGAATTTGAACCAATACGTA 714
QY 241 GlnAspGlySerLeuIleAspLeuCYsGlyAlaThrLeuLeuTrpArgThrAlaAspGly 260
| | | | |
Db 715 CAAGATGCTCTCTTATATGACCTTTGTGTGCAACTTGTGTGGCTGATCGACAGAAAGC 774
QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnIleAsnAlaAla 280
| | | | |
Db 775 CTTTCCATATCTCTTACGTGAAGCATTTAAAGCTTTTAAACAGAGATCAATGCAACT 834
QY 281 ArgProGlnCYsProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
| | | | |
Db 835 CGCGCGAGTGGCTCTGTAGGCTTCAACACATAGCTTCCCAAGTAAAGAGAGAGAT 894
QY 301 ValValGluGluLysGlnProTyrPalATyrLeuSerCYsGlyHisValHisGlyTyrHis 320
| | | | |
Db 895 GTTGTAGATGAAGAACACATGGGTATATCTAAACTCGGACATGCTCATGTATATCAT 954
QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluAlaGlyLysProMetCybArgThr 340
| | | | |
Db 955 AACTGGAGAAACAAGAAAGAAAGTGAACGCGCAAGATCGTGAATGCTTATGTGATGCT 1014
QY 341 ValGlyProTyrValProLeuTribLeuGlyCYSGluAlaGlyPheTyrValAspAlaGly 360
| | | | |
Db 1015 GTTGGCTCTCATGTTCTCTGTGGCTGGATGTGAAGCTGATTTTATGTGAGCCGCC 1074
QY 361 ProProThriAlaAlaPheThrProCYSGlyHisValCYsSerGluLysSerAlaLysTyr 380
| | | | |
Db 1075 CTTCCCAACCATGCGTTTGAAGCCGTGTGGCATGTGTGTTCAGAAAAAGCAACGCTTAC 1134
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCYsProPheCYs 400
| | | | |
Db 1135 TGGTCCCAATGCCACTTCTCTCATGTGACTCATATTTTCATGACCTGTCCCTTTGT 1194
QY 401 AlaThrGlnLeuValGlyGluGlnAncySileLysLeuIlePheGlnGlyProIleAsp 420
| | | | |
Db 1195 GCACATCACTGGCTGTGTGAACAAGCGTACATCAAGCTTATTTTCAAGAGACTTGAAC 1254

RESULT 5
US-09-843-905A-11
; Sequence 11, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; NUMBER OF SEQ ID NOS: 15
```

SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 1338
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (513)..(513)
OTHER INFORMATION: uncore
US-09-843-905A-11

Alignment Scores:
Pred. No.: 3,57e-172 Length: 1338
Score: 1642.00 Matches: 299
Percent Similarity: 81.38% Conservative: 42
Best Local Similarity: 71.36% Mismatches: 74
Query Match: 71.70% Gaps: 3

US-10-041-030-4 (1-420) x US-09-843-905A-11 (1-1338)

QY 3 SerProGlyGlnGluGlnHisCyAlaProAenLysGluProValIySryGlyGluLeu 22
DB 88 TCTCCCGGT-----GAGATGGCGCAGCCGAGAGAGCCATCAAGTATGATGAATC 141
QY 23 ValValLeuGlyTyraAngIyAlaLeuProAngIyAAspArgIyArGaTgLySerArg 42
DB 142 ATGCTCTGGGCTTACATGATGTTCTGGCAAGTGGGACAAGGCGCCGCGCAAGCGCGC 201
QY 43 PheAlaLeuTyraTyraProLyAlaAenGlyValIyProSerThrValHisValIle 62
DB 202 CTGGCACTGAGCGCGCGGTGCGACGCCAGCGGAGTGAAGCCAAAGCTCATGACCAACATC 261
QY 63 SerThrProGlnAlaSerLyAlaIleSerCyAlaGlyGlnHisSerIleSerTyThr 82
DB 262 TCCAGCCCGCTGCTCTCCAGGACGATGATACCGTGTGATGACGACGACATCTGTATAC 321
QY 83 LeuSerArgAngIyThrValValValGlyTyThrHisAAspLyAAspThrAAspMetPhe 102
DB 322 CTGTCCCGGAGCCATCGGTCAATGATGAGTATACATGATGATGATGATGATGATGATG 381
QY 103 GlnValIyArGserThrGlySerProIleAAspPheValIyThrAAspThrIleSerGly 122
DB 382 CAGATTGGCGCGCTCCACAGAAATGATGATGATGATGATGATGATGATGATGATGATG 441
QY 123 SerGlnAAspThrAAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAlaCyAsx 142
DB 442 GGA---GGGCGTCCCGAGGCGCTTCTGCCACAGACCATCTCCGCTATGCTGCGCGC 498
QY 143 IleValCyAAspArgAngIyProTyThrAlaArgIlePheAlaAAspIyPheAAspSer 162
DB 499 ATCTCTGTGACCGNCGCGCACCATTAATGCGCGCATCTATGCGCTGCTGATGCGC 558
QY 163 SerLyAAspIlePheLeuGlyGlyValAlaIleTyThrLyAAspProAAspGlyHisMet 182
DB 559 TCTAGCAACATCTCTCTGAGAGCGCGCGCAAAATGCGGAGCCCAAGATGCGCTGATG 618
QY 183 AspGlyLeuThrThrAAspGlyValIleValIleMetHisAAspArgIyGlyPheThrGlnGlu 202
DB 619 GATGAGACTGACCAATGAGTCTCTGATGATGATGATGATGATGATGATGATGATGATG 678
QY 203 SerGlnProGlyValIyArGaTgLyIleSerValCyAsGlyAAspValIyThrLeuArGlu 222
DB 679 TCAAGCCCGGGGTCTGTGGCGGAGATCTCGGTCTGTGGAGATGTATCAATGCGCGGAC 728
QY 223 ThrArgSerAlaGlnGlyArGlyValIyLeuValIyGlnSerGlnThrAAspValIleGlnAAsp 242
DB 723 AGCGGCTCAGCCGAGCGGAGAGCTGTGAGAAACAGATCAACGCTGCGAGAC 758
QY 243 GlySerLeuIleAAspLeuCyGlyAlaThrLeuLeuThrArGlnIleAAspGlyLeuPhe 262
DB 799 GGCTCTCTATGACCTGTGTGGGCGCACATCTGTGTGGGCGCACACCGGCGGCGGCTGCTG 858

QY 263 HisThrProThrGlnIyHisIleGlnAlaLeuArGlnGlnIleAAspAlaAAspPro 282
DB 859 CGGGCTCCACACTGAGAACATGAGAGCGCGCAGAGAGCAATGACAGCGCGCC 918
QY 283 GlnCyProValIyLeuAAspThrLeuAlaPheProSerIleAAspArgLyGlnValIle 302
DB 919 CAGTCCCGGCTGAGCTGACATCTGAGCTTCCCAAGCCAGACCGCTGCGGCGCAGAG 978
QY 303 GlnGluLyS---GlnProTrrAlaTyraLeuSerCyGlyHisValHisGlyTyThrHisAAsp 321
DB 979 CCGACAAACAGACAGCGCTGGGTATGATGCTGCGGAGGAGCGTCAATGATGATGATGATG 1038
QY 322 TrpGlyHisAAspSerAAspThrGlnAlaAenGlyValIyGlnCyProMetCyAAspThrVal 341
DB 1039 TGGGCTGGCGGAGAGCGGAGCGCGCGCAGAGAGCGAATGTCTCTGCGCGCTTGTG 1098
QY 342 GlyProTyraValIyProLeuThrLeuGlyCyGlyAlaGlyPheTyraValAAspAlaGlyPro 361
DB 1099 GGGCTTATGCTCTTATGCTTATGCTTGGCAGAGCGCGCTGCTGCTGATGATGATGATG 1158
QY 362 ProThrHisAlaPheThrProCyGlyHisValIySerGlyLySerAlaLyTyThr 381
DB 1159 CTTAGCCATGCTTTCACCTTGGCGGACAGTCTGCTGTGAGAGACATGCGCGTACTG 1218
QY 382 SerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCyAAspProPheCyAla 401
DB 1219 GCCACACACACACTGCGCCAGCGCACCATGCTTTCAGGCGCGCTGCTTTCGCGG 1278
QY 402 ThrGlnLeuValIyGlnGlnAAspCyIleLyLeuIlePheGlnIyProIleAAsp 420
DB 1279 GCGTGTATACCGGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 1335

RESULT 6
US-08-136-277-1
Sequence 1, Application US/08136277
Patent No. 5644045
GENERAL INFORMATION:
APPLICANT: MANDEL, Jean-Louis
APPLICANT: AUBOURG, Patrick
APPLICANT: MOSSER, Jean
APPLICANT: SARDE, Claude
TITLE OF INVENTION: X-LINKED ADRENOLEUKODYSTROPHY GENE AND
TITLE OF INVENTION: CORRESPONDING PROTEIN
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Young & Thompson
STREET: 745 South 23rd Street
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,277
FILING DATE: 15-OCT-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: B2272
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2750 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

```

/ TOPOLOGY: linear
/ MOLECULE TYPE: CDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 387..2624
/ US-08-136-277-1

Alignment Scores:
Pred. No.: 0.00763 length: 2750
Score: 120.50 Matches: 104
Percent Similarity: 33.94% Conservative: 63
Best Local Similarity: 21.14% Mismatches: 155
Query Match: 5.26% Indels: 171
DB: 1 Gaps: 26

US-10-041-030-4 (1-420) x US-08-136-277-1 (1-2750)
QY 2 PheSerProGlyngIngluHIsCyBaAlProAsnLysGluProValIysTyrglylu 21
DB 1167 TTCTCCGCC-----AAGTCCGGGAG 1187
QY 22 LeuValValLeuGlyTyraEnglyAlaLeuProAsnGlyAspArgGlyArgArgSer 41
DB 1188 CTGGGT-----GCAGAGGAGCGCGCGGAGAGGG 1217
QY 42 ArgPheAla---LeuTyryAspArgProLysAlaAsnGlyValIysProSerThrValHis 60
DB 1218 GAGCTGGCTACATGACATCGCGCTGCGTGGCCAC----- 1253
QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysAlsglyngInHIsSerIleSer 80
DB 1254 -----TCGAGAGAGATCGCTTCTATGGGGGCATGAGTGGAG 1292
QY 81 TyrThrLeu-----SerArgAsnGlnThrValVal 91
DB 1293 CTGGCCCTCTACAGCGCTCTCAACAGACCTGGCTCGCAGATCAACCTCATCTTCG 1352
QY 92 Glu-----TyrThrHisAspLysAspThrAspMetPheGlnValGlyArgSerThr 108
DB 1353 GAAAGCTGTGATGATTTATGTCGAGACGATTCCTATGAGATATGTGCGAGCGCTCG 1412
QY 109 GluSerProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThrAspGlu 128
DB 1413 GGCCTGCTCATGTGGCTGCTCCCATCATCATCGACCTGCTACTCAGAGTCAAGATGCA 1472
QY 129 AlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAspArgAsn 148
DB 1473 GAGCGCGTGAAGAGGACGCTTGAAGAAAGAGAGAGAGAGAGCTGTGAGCCAGCGCACA 1532
QY 149 GluProTyThr-----AlaArgIlePheAlaAlaGlyPheAspSerSerLysAsnIle 166
DB 1533 GAAAGCTTCACTATTCGCCGGAACCTCGACAGCGGCGCGAGATCCATTAGCGGATC 1592
QY 167 -----PheLeuGlyLysValAlaIleTyThrLysAsnProAspGlyHisMetAsp 183
DB 1593 ATGTCTCTGACACA--GGAGGTGACGAGCTGCTGCTACACAGCCCGGATGACAGAT 1651
QY 184 -----GlyLeu----- 185
DB 1652 GTTCCAGGATTTGAAGATGTTCAAGCTGTCACTTCAAGAGGCCAGGAGCTAGAGGA 1711
QY 186 ThrThrAsnGlyValLeu-----ValMetHisProArgGlyGlyPheThrGlu 201
DB 1712 CGCTCAGGCGGGGTCTGGGACCATAGCGCGGTGTGTCCTCGTGGAGGGCCCTGAA 1771
QY 202 GluSerGlnProGlyValIleThrArgGluIleSerValCysGlyAspValIleTyThrLeuArg 221
DB 1772 ATCCAGAGCCAGGT-----GGTGAATGGGAAACAGGGGATCATCTGCCA 1816
QY 222 Glu-ThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnVal---Ile 240
DB 1817 GAAATCCCATCGTCAAGCGCTCAGAGAGAGGTGTGTGGCCAGCTCAACATCAAGGAT 1876

```

```

QY 240 uGlnAspGly-----SerLeuI 246
DB 1877 GGAGGAGGACATGCATCTGCATCAAGGCCCAATGGCTGGCGGACAGAGCTCCCTGT 1936
QY 246 eAspLeuCyBglyAlaThrLeuLeuThrArgThrAlaAspGlyLeuPheHisThrProTh 266
DB 1937 CCGGATCCTGGGTGG-----CTCTGGCCCACTACGAGTGGTGTGCTCTCAACACCCC 1990
QY 266 rGlnIysHisIleGluAlaLeuArgGlnIleAsnAlaAlaAspProGlnCysProVa 286
DB 1991 ACCCAGGCGCATGTTCTACATCCGAC-----AGGCCCTACATCTCTGT 2035
QY 286 lGly-----LeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValAlaGluI 304
DB 2036 GGGCTCCCTCGGTGACAGAGTATATACCG-----GACTAGTGGAGGA 2080
QY 304 uLysGlnProThrAlaTy-----LeuSerCyBglyHisVa 316
DB 2081 CATGCAGAGAGAGGCTACTCGAGACAGACCTGAGACCATCTCGACGTGTGCACCT 2140
QY 316 lHis-----GlyTyThrHis-----AsnThrGlyHis---Ar 325
DB 2141 GCACCAATCTCTGCACGGGAGGAGGTGGAGGCTATGTGTGCTGAAGAGACGTCT 2200
QY 325 sSerAspThrGlu-AlaAsnGluArgGlyCysPrometCysArgThrValGlyProTyrv 345
DB 2201 GTGGGTGGCGAGAGAGAGATGGCATGGCCGATGTTCTACCAAGGCCCAAGTA 2260
QY 345 aProLeuThrLeuGlyCysBglyAlaGlyPheTyThrValAspAlaGlyPro----- 361
DB 2261 GCGCTCTCTGATGATGATCAACAGCG-----CGTGAGCATGGA 2299
QY 362 -----ProThr 364
DB 2300 CGTGAAGCAAGATCTTCCAGCGGCCCAAGAGCGGGCATTCCTGCTCTCAATCAC 2359
QY 364 lAsAlaPheThrProCysGly-----HisValCys-----SerG 375
DB 2360 CCACGGGCGCTCCCTGTGGAAATACACACACTTGTCAAGTTCATGGGAGAGGGCG 2419
QY 375 LuLysSerAlaLysTyThrSerGlnIlePro 385
DB 2420 CTGGAAGTTCGAGAGCTGACTGACGCTCC 2451

RESULT 7
US-08-479-403-1
/ Sequence 1, Application US/08479403
/ Patent No. 5869039
/ GENERAL INFORMATION:
/ APPLICANT: MANDEL, Jean-Louis
/ APPLICANT: AUBOURG, Patrick
/ APPLICANT: MOSSER, Jean
/ APPLICANT: SARDE, Claude
/ TITLE OF INVENTION: X-LINKED ADRENOLEUKODYSTROPHY GENE AND
/ NUMBER OF SEQUENCES: 23
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Young & Thompson
/ STREET: 745 South 23rd Street
/ CITY: Arlington
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22202
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/479,403
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:

```


NAME: PATCH, Andrew J.
 REGISTRATION NUMBER: 32,925
 REFERENCE/DOCKET NUMBER: B2272D1V
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-521-2297
 TELEFAX: 703-685-0573
 TELEX: 248425 EMBON
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2750 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 387..2624
 US-08-479-403-1

Alignment Scores:
 Pred. No.: 0.00763 Length: 2750
 Score: 120.50 Matches: 104
 Percent Similarity: 33.94% Conservative: 63
 Best Local Similarity: 21.14% Mismatches: 155
 Query Match: 5.26% Indels: 171
 DB: 2 Gaps: 26

US-10-041-030-4 (1-420) x US-08-479-403-1 (1-2750)
 QY 2 PheSerProGlyGlnGlnGlnHisCybAlaProAsnLysGluProValLeuTyArgLysGlu 21
 DB 1167 TTCTCGCC-----MAGTTCGGGAG 1187
 QY 22 LeuValValLeuGlyTyArgGlnAlaLeuProAsnGlyAspArgGlyArgLysSer 41
 DB 1188 CTGGTG-----GCAGAGAGGCGCGCGAGGCGG 1217
 QY 42 ArgPheAla--LeuTyArgLysProGlyValAsnGlyValLysProSerThrValHis 60
 DB 1218 GAGCTCGCTACATGACCTCGCTGCTGCGCCAAC----- 1253
 QY 61 ValIleSerThrProGlnIleSerLysValIleSerCysLysGlyGlnHisSerIleSer 80
 DB 1254 -----TCGAGAGAGATGCTCTTATGAGGCGCATGAGTGAG 1292
 QY 81 TyrThrLeu-----SerArgAsnGlnThrValVal 91
 DB 1293 CTGGCCCTGCTACAGGCTCTCTACAGAGACCTGCGCATCAACCTCATCTTCTG 1352
 QY 92 Gly-----TyrThrHisAspLysAspThrAspMetPheGlnValGlyArgSerThr 108
 DB 1353 GAAAGCTGTGATGATGATGCTGAGAGAGTTCCTCATGAGTATGCTGAGCGCTCG 1412
 QY 109 GluSerProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThrAspGlu 128
 DB 1413 GCGCTGCTATGCTGCTGCTGCCATCATCTGCACTGCACTCAAGATGAGATGCA 1472
 QY 129 AlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgGlyLeuAlaCysAspArgAsn 148
 DB 1473 GAGCGCGTGAAGAAGCAGCTTGGAAGAAAGAGAGAGAGAGCTGTGAGCGAGCCAC 1532
 QY 149 GluProTyThr-----AlaArgIlePheAlaAlaGlyPheAspSerSerLysAsnIle 166
 DB 1533 GAAGCCTTCACTATTCGCCCAACCTCTCGACAGCGCTGCAATGCAATTTGAGCGGATC 1592
 QY 167 -----PheGluGlyGlyValAlaLysThrLysAsnProAspGlyHisMetAsp 183
 DB 1593 ATGTCTGTACAA--GAGGTGACGAGAGCTGCTGCTACACAGCCCGGTGACAGAGAT 1651
 QY 184 -----GlyLeu----- 185
 DB 1652 GTTCAGAGTATTTGAAGATGTTCAGGCGCTGCACTTCAAGAGGCCAGGAGCTAGAGGA 1711

QY 186 ThrThrAsnGlyValLeu-----ValMetHisProAsnGlyGlyPheThrGlu 201
 DB 1712 CCTCAGCGGGAGTCTGGGACCATAGGCGGCTGTGCTGCTCGGTGAGAGGCCCTGA 1771
 QY 202 GluSerGlnProGlyValAlaTrpArgGlnIleSerValCysGlyAspValTyrThrLeuArg 221
 DB 1772 GATCGAGGCGCAGGT-----GGTGAATGTGAACAGGGATATCATCTGGCA 1816
 QY 222 Glu-ThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnVal--Le 240
 DB 1817 GAACATCCCATCATGTCACCGCTCAGAGAGGTGGTGGCCAGCTCATACATCGAGGT 1876
 QY 240 uGlnAspGly-----SerLeuI 246
 DB 1877 GAGAGAGGAGCATGCTCTGATATACAGAGCCCAATGAGCTGCGGAGAGCTCTGTGT 1936
 QY 246 eAspLeuCybGlyAlaThrLeuLeuTrpArgThrAlaAspGlyLeuPheHisThrProTh 266
 DB 1937 CCGGATCTCGGGTGG-----CTTGCGCCACGTACGCGTGGTGTCTTACAGCCGCC 1990
 QY 266 rGlnLysHisIleGlnAlaLeuArgGlnGlnIleAsnAlaAlaArgProGlnCysProVa 286
 DB 1991 ACCCGCAGCATATGTTCTATCATCCCGAG-----AGGCTTACATGTCTGT 2035
 QY 286 LGLY-----LeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValGlnGln 304
 DB 2036 GGGCTCTCGCTGACAGGTATGATCAACCG-----GACTCAGTGAAGGA 2080
 QY 304 uLysGlnProTrpAlaTyr-----LeuSerCysGlyHisAsp 316
 DB 2081 CATGCAAGAGAGGCTATCTCGAGAGAGAGCCTGGAAGCATCTGAGAGTGTGACCT 2140
 QY 316 His-----GlyTyrHis-----AspTrpGlyHis--Ar 325
 DB 2141 GCAACCATCTCTGACGCGGAGGAGGTGAGAGCTATGTGACTGGAAGAGCCTCT 2200
 QY 325 gSerAspThrGlu-AlaAsnGluArgGluCysProMetCysArgThrValGlyProTyV 345
 DB 2201 GTGCGGTGCGAGAGACAGAGATGCGATGCGCCGATGTTCTACACAGCGCCAACTA 2260
 QY 345 aLProLeuTrpLeuGlyCysGlnAlaGlyPheTyValAspAlaGlyPro----- 361
 DB 2261 CCGCTCTCGATGATGATGACCAAGC-----CGTAGACATGCA 2299
 QY 362 -----ProThr 364
 DB 2300 CTGGAAGGCAAGATCTTCCAGGCGGCAAGAGCGGCGCATTTGCTCTCCATCAC 2359
 QY 364 IeAlaPheThrProCysGly-----HisValCys-----SerG 375
 DB 2360 CCAAGCGGCTCTCTGTGAAATACCACACACACTTGCTTACAGTTGATGGGAGGCGG 2419
 QY 375 IuLysSerAlaLysTyrTrpSerGlnIlePro 385
 DB 2420 CTGGAAGTTGAGAAAGCTGACTCACTGCTGCC 2451

RESULT 8
 US-08-835-734-1
 Sequence 1, Application US/08835734
 Patent No. 6013769
 GENERAL INFORMATION:
 APPLICANT: MANDEL, Jean-Louis
 APPLICANT: AUBOURG, Patrick
 APPLICANT: MOSSER, Jean
 APPLICANT: SARDE, Claude
 TITLE OF INVENTION: X-LINKED ADRENOLEUKODYSTROPHY GENE AND
 CORRESPONDING PROTEIN
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Young & Thompson
 STREET: 745 South 23rd Street
 CITY: Arlington
 STATE: VA

```

COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,734
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/479,403
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: B2272D1V
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2750 base pairs
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 387..2624
US-08-835-734-1

Alignment Scores:
Pred. No.: 0.00763 Length: 2750
Score: 120.50 Matches: 104
Percent Similarity: 33.94% Conservative: 63
Best Local Similarity: 21.14% Mismatches: 155
Query Match: 5.26% Indels: 171
DB: 3 Gaps: 26

US-10-041-030-4 (1-420) x US-08-835-734-1 (1-2750)
QY 2 PheserProglYglnlgluHlslsCySaLaProAnlYgsluProVallyrTyrglygu 21
DB 1167 TTCTGCCCC-----AAGTTCGGGGAG 1187
QY 22 LeuValIValleuGlYTyraNglyAlaLeuProAnGlyAsPaRgLyArGaRgLySser 41
DB 1188 CTGGTC-----GCAGAGAGGCGCGCGGAGAGGG 1217
QY 42 ArgPheAla--LeuTyRlySaRgProLySaLaAnGlyVallyrProSerThrValHis 60
DB 1218 GAGCTCGCTACATGACATCGCTGCTGTCGCAAC----- 1253
QY 61 ValIleSerThrProGlnAlaSerLySaLaIleSerCyAlYsglyGlnHisSerIleSer 80
DB 1254 -----TCGAGAGATCGCTTCTATGCGGGCCATGAGTCGAG 1292
QY 81 TyrThrLeu-----SerArgAnGlnThrValValVal 91
DB 1293 CTGGCCCTGCTACAGGCTCTCTACAGACCTGCGCTCGCAATCACTCATCTTCTG 1352
QY 92 Gln-----TyrThrHisAsPlySaAsPThrAsPMeT PheGlnValGlyArGserThr 108
DB 1353 GAACGCTGTGATGTATGTGTGAGCAATTCCTCATGAGATGTGTGAGCGCTCG 1412
QY 109 GluSerProIleAsPheValValThrAsPThrIleSerGlySerGlnAnThrAsPglu 128
DB 1413 GGCCTGCTCATGTGCTGCTCCCATCATCATCGCCATCGCTACTACAGATCATGCA 1472
QY 129 AlaGlnIleThrGlnSerThrIleSerArgPheAlaCySaRglleValCySaAsPaRgaAn 148

```

```

DB 1473 GAGGCGGTGAAGAAGGACCTTGAGAAAAAGAGAGAGAGCTGTGTAGAGGAGCA 1532
QY 149 GlnProTyrrhr-----AlaArgIlePheAlaAlaGlyPheAsPserSerlySaAnIle 166
DB 1533 GAAGCTTCACATATTATGCCCGCAACTCTGACAGCGGTGTCAGATGCCATGAGCGGATC 1592
QY 167 -----PheLeuGlyGlnLysAlaIalYsThrPlySaAnProAsPglYHisMeLAsP 183
DB 1593 ATGTGCTGTACAA-GGAGGTGACGAGCTGGCTGTCTACACAGCCCGGTGCACAGANT 1651
QY 184 -----GlyLeu----- 185
DB 1652 GTTCAGGATATTGAAGATGTTCAAGCGTGCATTCAGAGGCCCGAGGAGCTAGAGA 1711
QY 186 ThrThrAnGlyValLeu-----ValMeThiAsProAnGlyGlyPheThrGlu 201
DB 1712 CGCTGAGCGGGGTCTGGAGACCATGAGCGCGTGTGTGTGCTGTGAGAGGCCCTTAA 1771
QY 202 GluSerGlnProGlyValThrArgGluIleSerValCySglyAsPValTyrrThrleuArg 221
DB 1772 GATCCGAGCCAGGT-----GATGATGTGAACAGGGGATCATCTGCGA 1816
QY 222 Gln-ThrArgSerAlaGlnAlaArgGlyLyLeuValGluSerGlnThrAsnVal---Le 240
DB 1817 GAACATCCCCCATCTGCACAGCCCTCAGAGAGGTGTGTGCGCACCTCAACATCAGGGT 1876
QY 240 uGlnAsPgly-----SerLeuI 246
DB 1877 GGAGGAGGCGATGATCTGCTCATCATCAGAGCCCAATGCTGCGGCAAGCTCCCTGTT 1936
QY 246 eAsPLeuCySglyAlaThrleuLeuThrPArgThrAlaAsPglyLeuPheHisThrProTh 266
DB 1937 CCGAATCTGGGTGG-----CTGTGGCCACGACGAGTGTGTCTCTAAGACCCCC 1990
QY 266 rGlnIleHisIleGlnAlaLeuArgGlnLylIleAsnIalAlaAlaProGlnCyAsProVa 286
DB 1991 ACCCGAGGCGATGTTCTACATCCGCAG-----AGCCCTACATGTCTGT 2035
QY 286 lGly-----LeuAnThrleuAlaPheProSerIleAsnArgLyGluValGluI 304
DB 2036 GGGCTCCCTGCGTGCATGACAGATGATCTACCG-----GACTCAGTGAAGA 2080
QY 304 uLySglnProThrAlaTyrr-----LeuSerCySglyHisVa 316
DB 2081 CATGCAAGAGAAAGGCTACTCTGAGACAGACCTGGAAGCCATCTGAGCTGTGCACCT 2140
QY 316 His-----GlyTyrrHis-----AsnTrpGlyHis---Ar 325
DB 2141 GCACCAATCTCGACGGGAGGAGGAGGTGGAGGCTATGTGTGACTGGAAGAGACTCT 2200
QY 325 gSerAsPThrGlu-AlaAnGlnuArgGluCyAsProMetCyAsArgThrValGlyProTyrv 345
DB 2201 GTTCGGGTGCGAGAGAGAGATGCGATGCGCCGCGATGTTCTACACAGGCCCAAGTA 2260
QY 345 aProLeuThrleuLysGlyuAlaGlyPheTyrrValAsPAlaGlyPro----- 361
DB 2261 GCGCTCTCGATGAATGACACAGCG-----CGTGAGCATGCA 2299
QY 362 -----ProThrH 364
DB 2300 CGTGAAGGCAAGATCTTCAAGGGGCGCAAGAGCGGGCATGCGCTGTCTCATCAC 2359
QY 364 leAlaPheThrProCySgly-----HisValCyS-----Serg 375
DB 2360 CCACCGGCGCTCCCTGTGGAATAACACACACACTTCTACAGTTGATGGAGGAGGCGG 2419
QY 375 LuLySerAlaIalYsTyrrTrpSerGlnIlePro 385
DB 2420 CTGGAAGTTGAGAGAGCTGACTGCTGCC 2451
RESULT 9
US-09-103-840A-2/c

```

Alignment Scores:	
Pred. No.:	7.4e+03
Score:	113.50
Percent Similarity:	34.02%
Best Local Similarity:	22.37%
Query Match:	4.96%
DB:	3
Length:	4403765
Matches:	98
Conservative:	51
Mismatches:	161
Indels:	129
Gaps:	21

[illegible]

Pred. No.: 7.43e+03 Length: 4411529
 Score: 113.50 Matches: 98
 Percent Similarity: 34.028 Conservative: 51
 Best Local Similarity: 22.37% Mismatches: 161
 Query Match: 4.96% Indels: 129
 DB: 3 Gaps: 21

US-10-041-030-4 (1-420) x US-09-103-840A-1 (1-4411529)

```

QY      4 ProGlyInglu-----GluHisValProAsnLysGluProValLysTyr 19
DB      1251307 CTTGGGCAACAAGGCGACGCGATGCAATGCGCCGCGACCATCTGCC----- 1251257

QY      20 GlyGluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArg 39
DB      1251256 GGTGCCCGCTGCATTTCTGATTACAC-----CGAGGTGGCGCT 1251218

QY      40 LysSerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysPro----- 56
DB      1251217 -----GCGCTATCTGCGCGCGCGGATCCAGATCGCGCGATCGGGATGA 1251170

QY      57 -----SerThrValHisValLleSerThrPro 65
DB      1251169 CTTGCTGATCGGATGATGCTGCTGGAAACAGAAACCTGTGCAGTTCGCGCTGGCCC 1251110

QY      66 GluAlaSerLysAlaLleSerCysLysGlyGlnHisSerLle-----Ser 80
DB      1251109 GAATCGCGCTGCTGTGACGCGCTCGGAGGCCACCCAGTCGATGCCGCGATATGGA 1251050

QY      81 TyrThrLysSerArg-AsnGlnThrValValGlu----- 92
DB      1251049 CACGATTCCTGACGGGGGATACCGACGCGGTCCAAACAGATCTGCCAGGCGCGGTCC 1250990

QY      93 -----TyrThrHisAsp-LysAspThrAspMet---PheGlnValGlyArgSerThr 109
DB      1250989 CGCCGACCTTCGCCACGACACGACGCGGATTCACAGCATGTATCGCGCATGTCGG 1250930

QY      109 LysSerProLysAspPheValValThrAspThrLleSerGlySerGlnAsnThrAspGlu 128
DB      1250929 CAGCGCGGTCCCATTTACCGACGACCCGATGTCGCCGAGCAGAGCGGT 1250870

QY      129 -----AlaGlnLleThrGlnSerThr- 135
DB      1250869 ATTGCGCTCTGGAACTCACTCGGCGCTTCATGCTATCGGACATCAAGCGCTGACGA 1250810

QY      136 -----LleSerArgPheAlaCysArgLleValCysAspArgAsnGluProTyrThr 153
DB      1250809 CCAGGCGGCTACGCGGCTGTGTGTCAGGCGATGTATTGGCAACTCACC----- 1250757

QY      153 LArgLlePheAlaLaglyPheAspSerSerLysAsnLlePheLeuGly-----GluL 171
DB      1250756 --GAACTCG--GCGCGCGGGGTCTGTCTCAAGCAACCGGATCAACCGGAGCCAAAGAC 1250700

QY      171 YsAlaLalysTrpLysAsnProAspGlyHisMetAspGlyLeuThrThrAsnGlyVal 191
DB      1250699 CGTTGGCGGAATGCTGA---CCCGCGCGCCAAAGGCGACACGTC---GTGTCCGATTGG 1250646

QY      191 euValMetHisProArgGlyGlyPheThrGlnGlnSerGlnProGlyValTyrArgGly 211
DB      1250645 TCGCGGTGCGACGATCAACCGGCTCGCGACAG---CAGCGCGCGCTCTACGACCAAC 1250589

QY      211 leSerValCysGlyAspValTyrThrLeuArgGluThrArgSerAlaGlnGlnArgGly 231
DB      1250588 TG-----ATGCGCGCGCGCGCGGATTGGAGCCCATGGCGAG 1250553

QY      231 YsLeuValGluSerGluThrAsnValLeuGlnAspGlySer-----LeuLleAspLeuc 249
DB      1250552 GCGAGTCCAGAGATTCGAATTCACCGGTGAGAGACAGCCAGCTGTGCTGTGCGACACC 1250493

QY      249 YsGlyAlaThrLeuLeuTyrArgThrAlaAspGlyLeuPheHisThrProThrGlnLys 269
DB      1250492 GCGGCGCGGAA-----CGCTCGCGCGAGCGCGCGCTGTGCGACTGGCGCTGCAACTGC 1250442
  
```

```

QY      269 isLleGluAlaLeuArgGlnGluIleAsnAlaAlaArgProGlnCysProValGlyLeu 289
DB      1250441 ATCAGAGAGACTCATTCAGACGACCGGACACTGCGCGGTACTCTCGAGCCATTCG 1250382

QY      289 snThrLeuAlaPheProSerLle-AsnArgLysGluValAlaGluGlnLysGlnProTrp 308
DB      1250381 AGACTTCTGTCAGCGCGTCTGCGCGACGAGAA----- 1250349

QY      309 AlaTyrLysSerCysGlyHisValHisGlyTyrHisAsnTrpGlyHisArgSerAspThr 328
DB      1250348 -----CAGCGTGTGCTGCACCGC----- 1250331

QY      329 GluAlaAsnGluArgGluCysProMetCysArgThrValGlyProTyrValProLeuTrp 348
DB      1250330 ---TTTGGCCAGAGGCTTGCCTTCCCGGCGGTGTATCC----- 1250289

QY      349 LeuGlyCysGluAlaGlyPheTyrValAspAlaGlyProProThrHisAlaPheThrPro 368
DB      1250288 -----GGACCGCTTACACCGAGGTGACGAGG 1250262

QY      369 CysGlyHisValCysSerGluLysSerAlaLysTyrTrpSerGlnIle 384
DB      1250261 CGTTGACGCTGCGGACCGGCGGCGCGGTGATCTTGTGCGAGATC 1250214

RESULT 11
US-09-252-991A-14397/C
; Sequence 14397, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT FILING DATE: US/09/252, 991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14397
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14397

Alignment Scores:
Pred. No.: 0.0591 Length: 1590
Score: 109.00 Matches: 107
Percent Similarity: 30.08% Conservative: 47
Best Local Similarity: 20.90% Mismatches: 177
Query Match: 4.76% Indels: 182
DB: 4 Gaps: 25

US-10-041-030-4 (1-420) x US-09-252-991A-14397 (1-1590)

QY      24 ValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArg-----GlyArgArgLys 40
DB      1477 CTGCTCGCGCGGATGCGGTGCTCCA---GGCGAAAGCGACGTCGCGGTGCGGAAAG 1421

QY      41 SerArgPheAlaLeuTyrLysArg-----ProLysAlaAsnGlyValLysProSer 57
DB      1420 CTTCGTGCGGCGCTCATCATCATCGTTGCGAATTCCACTGCGCACATCGATCGATGACCGA 1361

QY      58 ThrValHisValLleSerThrProGlnAlaSerLysAlaLleSerCysLysGlyGlnHis 77
DB      1360 ACTGCGCGCGCTTCGAGATCTCCGCTGCTGCGGCGGCGCTGATCACCGGAT 1301

QY      78 -----SerLysSerTyrThrLysSerArg----- 85
DB      1300 CCACCATGACTGCTGCGCCCGAGAGATCCGTGCGCGAGCATATCGCTGTCCGACAC 1241

QY      86 -----Asn 86
  
```

Db 1240 CTGGCCGGGGCCACTGCTGACCAACCCCGCCGACGAGCGCGCGGCGGAAAC 1181
 QY 87 GlnThrValValAluValuTyrThrHisAspLysAspThrAspMetPheGlnValGlyArg 106
 Db 1180 CGATCGGGGGCGGACCGGATCGGCCCAACGCGGGAACGCTGCGGCGCTGACTCGG 1121
 QY 107 SerThrGluSerProIleAspPheValValThrAspThrIleSerGlySerGln-AsnTh 126
 Db 1120 ACCGCG-AGCGGTCCGATTGATTCGTATAC-----GAAATACGAAAAAGAGAAACCA 1068
 QY 126 rAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValGlyAsn 146
 Db 1067 TGAAGCAGCAGCGTCAACGAGCGCGGTATCCCGCG----- 1031
 QY 146 rArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPheAspSerSerLysAsnI 166
 Db 1030 -----GGTCACACCCCAATTCAGCGCG 1008
 QY 166 ePheIleGlyGluValAlaAlaLysTrpLysAsnProAspGlyHisMetAspGlyLeuTh 186
 Db 1007 ATTTCCTCGGTGAACCTGAGGA-----AACCCACACGCGTCATCTC----- 968
 QY 186 rThrAsnGlyValLeuValMetHisProArgGlyGlyPheThrGluGluSerGlnProG 206
 Db 967 -----CAACCTGG 960
 QY 206 yValTrpArgGluIleSerValCysGlyAspValTyrThrLeuArgGluThrArgSerAl 226
 Db 959 TGGCGGACCGCGGTGCGG-----GCTGTGGGTATGCGGACGCGCGGAGAACACTCGG 993
 QY 226 eGlnGlnAspGly-LysLeuValGluSerGluThrAsnValLeuGlnAsp-----G 243
 Db 902 TGCACATCGAGAGAAATG-----GCGGTCAACGAGGCGGCAAGATGCTCCGCGG 849
 QY 243 ySerLeuIleAspLeuCysGly-----A 251
 Db 848 GCGCGGTGCGGTGATGCGGGATGCGCGGATTCACACGCGCGCGCGGAGAGTGG 789
 QY 251 yThrLeuLeuTrpArgThr--AlaAspGlyLeuPheHisThrPro----- 265
 Db 788 CCAAGCGGTGAGCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 729
 QY 266 -----ThrGlnLysHisIleGlnIleValLeuArgGlnGluIle----- 277
 Db 728 CTTCCAGCGCGGACGAGACCGCGGACATTCGCGAGCGGTGCGGACGCGGACGCTGC 669
 QY 278 -----AsnAlaAlaArgProGlnCysArgPro 286
 Db 668 CGGTGATGCTTACACCAACCGCGGATTCACCGGAAACGATCCACCGCGGACATCTGG 609
 QY 286 aGlyLeu-----AsnThrLeuAlaPhe-----ProSerIleAsnArgLys 300
 Db 608 TTTCCCTGGCCGACATGCGAGAACATGCTGCTTCAAGACAGCTCCGCGGACCGCC 549
 QY 300 yValValAluGluLysGlnPro-TyrAlaTyrLeuSerCysGlyHisValHisGlyTyr 319
 Db 548 GCTTCATCGACGATCAACCAACAGTGGGAGCGCTTCATCTCTTCG--CCGGGCTCG 492
 QY 320 HisAsnTrpGlyHisArgSer-----AspThrGluAlaAsnGluArgGluCysPro 336
 Db 491 ACAGCTGTGCTCTGGAAGCGTGGCGGTGGGTCGCAAGGCTGATCTCGGAAATGTCCA 432
 QY 337 MetCys----- 338
 Db 431 AGGTGTCCGGAAGAGGAGACATCTTCGCGCTGCGGACGCGGCGGCTTGGCGG 372
 QY 339 -----ArgThrValGlyProTyrValProLeu 347
 Db 371 AGCGGATGCGGATCTAGAGTGGCTGATGCGGATCTCCATCTCGACGCGCGCTCCGAC 312
 QY 348 TrpLeuGlyCysGluAlaGlyPheTyrValAlaAspAlaGlyPro----- 361

Db 311 TGG-----TGCAGTGCATCAAGCTCTGCGAGCAACTGCGGTCGCGGACGCGCTCA 258
 QY 362 ProThrHisAlaPheThrProCysGlyHisValCysSerGlnLysSerAlaLysTyrTrp 381
 Db 257 CCGGTCCACGCGGCGCTGCGGCTGCGGCGCGCGCGGATCGAGACGCTCGAGCGGATGAG 198
 QY 382 SerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaLys-----Pro 398
 Db 197 CCAAGG---CCCTGGCGACCGCGCGCGCTGCGGACATCGGCTTGTAGCGGACCGCC 141
 QY 399 PheCysAlaThrGlnLeuValGlyGln 408
 Db 140 GCGCTGCGCGCTGCTGCTGCGGAGCGG 111
 RESULT 12
 US-09-252-991A-14280
 ; Sequence 14280, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 14280
 ; LENGTH: 3801
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-14280
 Alignment Scores:
 Pred. No.: 0.244 Length: 3801
 Score: 109.00 Matches: 107
 Percent Similarity: 30.08% Conservative: 47
 Best Local Similarity: 20.90% Mismatches: 177
 Query Match: 4.76% Indels: 182
 DB: 4 Gaps: 25
 US-10-041-030-4 (1-420) x US-09-252-991A-14280 (1-3801)
 QY 24 ValLeuGlyTyrAsnGlyValAlaLeuProAsnGlyAspArg-----GlyArgArgLys 40
 Db 802 CTGCTCGCGCGGATGCGGTGCTCA---GGCGAAAGCCAGCTCGGCTGCGGAAAG 858
 QY 41 SerArgPheAlaLeuTyrLysArg-----ProLysAlaAsnGlyValLysProSer 57
 Db 859 CTTGCGGCGGCTCATCATGATCGGATTCGGAATTCACCTGCACATGATCGGTACCGCA 918
 QY 58 ThrValHisValIleSerThrProGlnAlaSerLysAlaIleSerCysGlyGlnHis 77
 Db 919 ACTCGGCGCGGCTGCGGATCTCCCGTGTCCGCGGCGGCGGCTGATCACCGGAT 978
 QY 78 -----SerIleSerThrLeuSerArg----- 85
 Db 979 CCAACGATCTGCTGCGCGCGGAGATCTCTGGCGCGAAGCTATCGGCTGTCCGACAC 1038
 QY 86 -----Asn 86
 Db 1039 CTGGCCGGGCGCACTGCTGACCAACCCCGCCGACGAGGACGCGCGGCGGAAAC 1098
 QY 87 GlnThrValValAluValuTyrThrHisAspLysAspThrAspMetPheGlnValGlyArg 106
 Db 1099 CGATCGGCGGCTACCGGATGCGCCACGAGCGGAGACGCTGCGGCTGACTCGG 1158
 QY 107 SerThrGluSerProIleAspPheValValThrAspThrIleSerGlySerGln-AsnTh 126
 Db 1159 ACCGCG-AGCGGTCCGATTGATTCGTATAC-----GAAATACGAAAAAGAGAAACCA 1211

```

QY      126 rAaPGLuaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAs 146
      ::::|||||:
Db      1212 TGAGCAAGACAGCTCACTGAGCGGCTATTCCCGC----- 1248
QY      146 pArgAnGluProTyrrThrAlaArgIlePheAlaAlaGlyPheAerSerIleAsnI 166
      ::::
Db      1249 -----GGTCACCAACCCCAATTCACACGCCG 1271
QY      166 ePheLeuGlyGlyIuValAlaAlaLysTrpLysAsnProAerGlyIleMetAerGlyLeuTh 186
      :|||||:
Db      1272 ATTTCCTCGGTGAACCTGAGAGA-----AACCCACACGGTCACTCTC----- 1311
QY      186 rThraGnGlyValLeuValMetHisProArgGlyGlyPheThrGlnGluSerGlnProG 206
      :|||:
Db      1312 -----CAACTCG 1319
QY      206 yValTrpArgGluIleSerValCysGlyAerValTyrrThrLeuArgGluThrArgSerAl 226
      :|||:
Db      1320 TGCGGAGCGCGGTGCCG-----GCTGTGTGTATGCGGACGCTCGGCGAAGACACTCGC 1376
QY      226 aGlnGlnArgGly-LysLeuValGluSerGluThrAsnValLeuGlnAer-----G 243
      :|||||:
Db      1377 TGACCATCGAGGAGAAAGATG-----GCCGTACGAGAGGTCCCAAGATGCCCTCGCG 1430
QY      243 ySerIleuIleAerLeuCysGly-----A 251
      :|||:
Db      1431 GCCCGGTGCCGTGATCTCGCGGATCCGCGAGTTCACACGCCCGCGGCGAAGATGG 1490
QY      251 lArhTrleuLeuTrpArgThr--AlaAerGlyLeuPheHisThrPro----- 265
      :|||:
Db      1491 CCAAGCGGTGAGCGGCTCGCGCTGATCGGATCGTGTGATGCCCGCGCTGTCTATT 1550
QY      266 -----ThrGlnYHisIleGluAlaLeuArgGlnIle----- 277
      :|||:
Db      1551 CCTCAAGCGGACAGACCGCGGACATTCGCGAGCGTGGCCAGCGGACCGACTGTC 1610
QY      278 -----AsnAlaIleArgProGlnCysPro 286
      :|||:
Db      1611 CGGTGATGCTACAAACACCGCGGATCTACCGGACACACTACCGCGGACATCTCG 1670
QY      286 aGlyLeu-----AerThrLeuAlaPhe-----ProSerIleAerArgLys 300
      :|||:
Db      1671 TTTCCTCGCGCACTGCGAGAACATCTCTCTCAAGGACAGCTCCGCGACACCGCG 1730
QY      300 lValValGlnGluLysGlnPro-ThrAlaTyrrLeuSerCysGlyHisValHisGlyTy 319
      :|||:
Db      1731 GCTTCATCGACGTACGCAACCAAGTGGCGAGCGCTTCATCTCTTCG---CCGGAGCTCG 1787
QY      320 HisAerTrpGlyHisArgSer-----AerThrGlnAlaAsnGluArgLysCysPro 336
      :|||:
Db      1788 ACGAGTGTCTCTGAAAGCGTGGCGGTGGTCCCAAGGCTCGATCTCGGAAATGTCA 1847
QY      337 MetCys----- 338
      :|||:
Db      1848 ACGTGTTCGCGAGAGAGCGAGACCATCTTCGCTGCGCGACGCGGCGCTTCGCG 1907
QY      339 -----ArgThnValGlyProTyrrValProLeu 347
      :|||:
Db      1908 ACGGATGCCGATCTACGAGTGGCTGATGCCGATCTCTCATCTCGACGCCCTCCGAG 1967
QY      348 TrpLeuGlyCysGlnAlaGlyPheTyrrValAerAlaGlyPro----- 361
      :|||:
Db      1968 TGG-----TGCAGTGCATCAAGACTCTGAGACAACTCGCGGTGGCGGACGCGCTCA 2021
QY      362 ProThrHisAlaPheThrProCysGlyHisValCysSerGlnLysSerAlaLysTyrr 381
      :|||:
Db      2022 CCGCTCCACCGCGCTGGCGCTGGCGGCGCGGATCGGACGACGCTCGAGCATATG 2081
QY      382 SerGlnIleProLeuProHisGlyHisAlaPheHisAlaAlaCys-----Pro 398
      :|||:
Db      2082 CCAAG---CCTGCGAGCGCGCGCGCTCGCGGACATCGGCTGTGAGCGGACGCGCC 2138

```

```

QY      399 PheCysAlaThrGlnLeuValGlyGln 408
      :|||:
Db      2139 GCCCTCGCGCGCTCGCTCCCGGAGCGG 2168

RESULT 13
US-09-799-451-571
; Sequence 571, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Duntui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/799, 451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pf_FL_genes Version 2.0
; SEQ ID NO 571
; LENGTH: 2240
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1779)
; US-09-799-451-571

Alignment Scores:
Pred. No.: 0.196 Length: 2240
Score: 106.50 Matches: 68
Percent Similarity: 36.53% Conservative: 31
Best Local Similarity: 25.09% Mismatches: 107
Query Match: 4.65% Indels: 65
DB: 4 Gaps: 16

US-10-041-030-4 (1-420) x US-09-799-451-571 (1-2240)
QY      53 GlnValLysProSerThrValHisValIleSerThrProGlnAlaSerLysAlaIleSer 72
      :|||:
Db      649 GGAGGCCAGCCTTCACAGACAGCCAGGCGCTGACGAGCACTGCCACAGGCTTTTTC 708
QY      73 CysLysGlyGlnHisSerIleSerTyrrThrIleuSerArgAnGlnThnValValGlu 92
      :|||:
Db      709 -----CACACACAGCCCGCTCTTGCGCGGACCGGTAGACTTGCGACAGA 756
QY      93 -----TyrrThrHisAerLysAerThrAerMetPheGlnValGlyArg 106
      :|||:
Db      757 AGAATTGATCAAACTGTGTCAAACTATCAAGGTACACTGTGTGAGATCTGTGCGC 816
QY      107 SerThrGlnSerProIle--AerPheValValThrAerThrIleSerGlySerGlnAsn 125
      :|||:
Db      817 CAGGACGAGTCACTTCTCAAGACAGCTGTGACA-----CAAGGAGAGGAAAGG 867
QY      126 ThrAerGlnAlaGlnIleThrGlnSerThrIleSerArgPheAlaCys----- 141
      :|||:
Db      868 GAGACACAGCCAGCTGTGAGATCTTGTGTCCAGCTG---TGCCCTCAAGCGGCGC 924

```



```

QY 142 -----ArgIIeValCyAaPrArgAnGlnProTyrThrAlaArgIle 155
DB 925 CAGGCAATTGACCTGGGCGAGCTTCTGCAAAAGAAAGCCCTGCGCTGGCGGCG 984
QY 156 Phe-----AlaAlaGlyPheAaPseSerIyAaenIlePheLeuGly 169
DB 985 CTGCTTCAGAGAGAGACCCCGGACGCGTTCTGAGAGAGTGCAGAGAACTTGTGTGGG 1044
QY 170 -----GlnIyAaAlaAlaIyTPrIyAaen-----ProAaPglYhiSeAaP 183
DB 1045 CTTCAGAACAGAGAAACCTGCTGCTGCTGAGCCCAACATCAAGATGGTGAT---GAG 1101
QY 184 GlnIyLeuThrThraAaGlyValIleuValMetIleProAaGlyGlyPheThrGluGluSer 203
DB 1102 AGTGGCGGTGTGAATGAGCTTGATAGCAAAATGACTCTTGAAGAGC-----AGT 1149
QY 204 GlnProGlyValIyTPrArgGluIleSerValCyAaGly-----AaPValIyThrIleu 220
DB 1150 CCGCCAGATATTGGAGACAAATTTCAGAAATGTTCTATGTGGAGAGACGTCCTCTCTG 1209
QY 221 ArgGluThrArgSerAlaGlnIyAaGlyIyValIleuValGluSerGluThrAaenValIleu 240
DB 1210 GCCGTGGGCGGACGAGACCTGACAGAGAGTCTCCCAAGACATCTGGAACAGCTCTTA 1269
QY 241 GlnAaPglYSerIleuIleAaPleuCyAaIyAlaThrIleuLeuTPrArgThrIlaAaPglY 260
DB 1270 -----GGCCAGCTG-----GGCCAGACCTGCGGTGGCGGAG----- 1302
QY 261 LeuPheIleThrProThrGlnIyAaIleGluAlaIleuAaGlnIyIleAaenAlaIa 280
DB 1303 TTCCTGTGCCACCTCTGTAGCAGCATCTG----- 1332
QY 281 ArgProGlnCyAaProValGlyIleuAaenThrIleu 291
DB 1333 ---GCAAGTGTCTGTGAGTGAAGTTCCTC 1362

RESULT 14
US-09-103-840A-2
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Alignment Scores:
Pred. No.: 4.95e+04 Length: 4403765
Score: 106.00 Matches: 85
Percent Similarity: 35.90% Conservative: 55
Best Local Similarity: 21.79% Mismatches: 140
Query Match: 4.63% Indels: 112
DB: Gaps: 19

US-10-041-030-4 (1-420) x US-09-103-840A-2 (1-4403765)
QY 10 CyAaIaProAaenIySerIyValIyGlyGluLeuValIleuGlyTyr----- 27

```

```

DB 3782406 TGTGGCGGAAATCAAGGCGCTTACATTGGAGAGATCAACCGCGCTGGAGGGGCGCT 3782465
QY 28 -----AaenGlyAlaIleuProAaen 33
DB 3782466 CAACGAGATTGATATCAAGACGCGTGTGTAAACCAGCCAAATTAGGAGATCTCTTC 3782525
QY 34 GlnAaPArgGlyIyAaGlySerAaPheAlaIleuTyrIyAaArg-----ProIyA 50
DB 3782526 GAGCTGGCGCGGACGAGCTGAGCTGCTGAGGAGCAAGACGAGTATGCGCTCA 3782585
QY 51 AlaAaenGlyValIyAaPseSerThrValIleValIleSerThrProGlnIaIleSerIyAaIa 70
DB 3782586 CTATATGGCTTGGCGCCAGGAGATATTCAGACCGCGCTCCAGGCTTGAAGAAACCGCGG 3782645
QY 71 IleSerCyAaIyGln-----His-SerIleSerTyrThrIleuSerArgAaenIlnTh 88
DB 3782646 CCGTGGCGTCAAGGTGGGTCTTCAAAACAAACAGCGTGTGATGAGCGCGGAGCGCT 3782705
QY 88 rValIyAlaIyGluTyrThrIleAaPlyAaPThrAaPheGlnValIyIyAaGserTh 108
DB 3782706 CCTTACAGTGGCGCGCTGTGACAGACCTC----- 3782733
QY 108 rGluSerProIleAaPheValIlnThAaPThrIleSerGlySerGlnAaenThrAaPgl 128
DB 3782734 -----GTCAGCTGTGTGTATGCGAGATGATGAGAGCTGCAACCTGACCC 3782783
QY 128 uAlaGlnIleThrGlnSerThrIleSerAaPheAlaCyAaGlyIleValIyAaPArgAa 148
DB 3782784 G----- 3782784
QY 148 nGluProGlyTyrAlaArgIlePheAlaIleAaPseSerIyAaenIlePheIe 168
DB 3782785 -CGGCGTATCAAGGATGCGGAGCGGACCTCGCGTCTGCAACAGTCATGCTGTCTT 3782843
QY 168 uGlnIyAaIyAlaIyTPrIyAaAaPProAaPArgIyIleMetAaPglYleuThrThra 188
DB 3782844 C---GACGACATCGCGACTG-----GTTGAGGCGCCACGCTGCC 3782882
QY 188 nGlyVal--LeuValMetIleProAaGlyIyPheThrGluGluSerGlnProGlyAa 207
DB 3782883 GGGCATGCGCGGATACCTGTGACCGTTCGCGAACAATCGGCAC-----GGCGT 3782933
QY 207 ITrPrArgGluIleSerValCyAaIyAaPValIyThrIleuAaGluThrArgSerAlaI 227
DB 3782934 CATTGCGATTTGTCTCAGGCTTGAGACATCTTGACCG--CGGCGACCATGACCGAAG 3782992
QY 227 nGlnAaGlyIyIleuValIleuSerGluThrAaenValIleuGln----- 241
DB 3782993 TGACGAGCCGACATCGCGGACAGGAGCGCTCATTTGTTACTTCATGACAGCGCGCA 3783052
QY 242 -----AaPglYSerIleu-----IleAaPleuCyAaIy----- 250
DB 3783053 GTGCACTTAGATGAGTGAAGCTTAACTAACCGTGGCGGAGAGTAAAGTCTTGCAATCC 3783112
QY 251 -AlaThrIleuLeuTPrArgThrIlaAaPglYleuPheIleThrProThrGlnIyAaIle 270
DB 3783113 ACAAGCGCGCGTATGCGGTTGCA-----GTGTTGCGATGACCAACCGCGCGAG 3783166
QY 270 eGluAlaIleuAaGlnIyIle-----AaenAlaIaArgProGlnIy 284
DB 3783167 TGAATCGCGCGACAAAG-ATGTGCGCGTGTGCTGTGCAATTCAGAGACTTGAATG 3783225
QY 284 aProValIyIleuAaenThrIleuAlaPhePro--SerIleAaenArgIyGluValIaI 303
DB 3783226 C-----AATATGCTGTCTCTTCGAGATGCGTGGCGGTGCAAAATCCCAAC 3783273
QY 303 uGlnIySerIleProTPrAlaIyTyr-----LeuSerCyAaIyIleValIa 316
DB 3783274 GAGCATGCGGCGCTGCGTATTAAGATTCACCAATGCGCATTCGAGCATGAGGCGATCA 3783333
QY 316 IHisGlyTyrHisAaenTPrGlyIleAaGserAaPThrGlnIleAaenGluAaGlyCyAa-- 335

```

```

Db      3783334 TCATAGTAC-----TGCGACGACATATCAGATGGCGAGTGTATTAATATGCGCATGG 3783387
QY      336 -----PrometCysArgThrVal 341
Db      3783388 TGCCACTTATCCGCCGACGAGAACCGTC 3783415
RESULT 15
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6284328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 4.97e+04 Length: 4411529
Score: 106.00 Matches: 85
Percent Similarity: 35.90% Conservative: 55
Best Local Similarity: 21.79% Mismatches: 140
Query Match: 4.63% Indels: 112
DB: 3 Gaps: 19

US-10-041-030-4 (1-420) x US-09-103-840A-1 (1-4411529)
QY      10 CysAlaProAsnLysGluProValLysTrpGlyGluLeuValLeuGlyTrp----- 27
Db      3790226 TGTGGCGCAATACACGGGCTTACATTGGAGAGATCGAACCGCTGGGGGTGGTGGCT 3790285
QY      28 -----AsnGlyAlaLeuProAsn 33
Db      3790286 CAACGATTGACTATGACGAGCGCTTCTGAAACCCAGCCATTAAGGAGTCTCTTC 3790345
QY      34 GlyAspArgGlyArgTrpLysSerArgPheAlaLeuTrpLysArg-----ProLys 50
Db      3790346 GAGCGTGGCGGCGAGCTGAGCTGCTTCAAGGCAAGAGAGAGCTAGTGGCGCTCGA 3790405
QY      51 AlaAsnGlyValLysProSerThrValHisValHisSerThrProGlnAlaSerLysVal 70
Db      3790406 CTACATGGGCTTCGCGCGAGATATCCAGCGCGCTTCAAGCCCTTGAAGAAGCCCGGCG 3790465
QY      71 HisSerCysLysGlyGln-----His-SerLysSerTrpThrLysSerArgAsnGln 88
Db      3790466 CCGTGCGCTCAAGAGTGGCTTCTCAACAACAACGCTGTTGTTGACGCCCGCAGCCT 3790525
QY      88 rValValValGluTrpThrHisAspLysAspThrAspMetPheGlnValGlyArgSerThr 108
Db      3790526 CTTTCAGTGGCGCGCTCTGACAGACTC----- 3790553
QY      108 rGluSerProLysPheAspPheValValThrAspThrLysSerGlySerGlnAsnThrAspG 128
Db      3790554 -----GTCGACGCTGCTGCTGATTCGACAGATGATCGAGGCTGACAGCCTGACCC 3790603
QY      128 uAlaGlnIleThrGlnSerThrLysSerArgPheAlaCysArgIleValCysAspArgAs 148
Db      3790604 G----- 3790604
QY      148 nGluProTrpThrAlaArgIlePheAlaAlaGlyPheAspSerSerLysAsnIlePheLe 168

```

```

Db      3790605 -CGGCGCTATCAAGGATCGCGGAAGCCCTCGGCGTCTCGACAACGTCATGCTGTCTT 3790663
QY      168 uGlyGluValAlaAlaLysTrpLysAsnProAspGlyHisMetAspGlyLeuThrAs 188
Db      3790664 C--GACGACATTCGCGCACTG-----GTGAGGGCGGACGGTGGCC 3790702
QY      188 nGlyVal--LeuValMetHisProArgGlyGlyPheThrGluGluSerGlnProGlyVal 207
Db      3790703 GGGCATGGCGCGGATACCTCGTGAACCGTTCCGGACAACCTGGCAGC-----GGCGT 3790753
QY      207 LTrpArgGluLysSerValCysGlyAspValTrpThrLeuArgGluThrArgSerAlaG 227
Db      3790754 CGTTGCGATTTGTTCAGCCCTTGGAGCATCTTGAGCCG-GCGGAGCAATGACCGAACG 3790812
QY      227 nGlnArgGlyLysLeuValGluSerGluThrAsnValLeuGln----- 241
Db      3790813 TGACGAGCGGACATCGCGCAGAGAGCGCTCATGTGTACTCATGACCCAGCCGCA 3790872
QY      242 -----AspGlySerLeu-----IleAspLeuCysGly----- 250
Db      3790873 GTGCATTAAGATGGCAGCTTAACTAACCCGTCGCGAGCAGATTAAGTCTTGGCAATCC 3790932
QY      251 -AlaThrLeuLeuTrpArgThrAlaAspGlyLeuPheHisThrProThrGlnLysHisI 270
Db      3790933 ACAACGGCGCTATGCGCGTTGCGA-----GTGTGCGATAGACCAACCCGCGCAG 3790986
QY      270 eGluAlaLeuArgGlnGluIle-----AsnAlaAlaArgProGlnCy 284
Db      3790987 TGATTCGCGCGGACAGG-ArgTGGCGCTGTGCTTCCCAATGGCGGACGCTGAATG 3791045
QY      284 sProValGlyLeuAsnThrLeuAlaPhePro--SerLysAsnGlyGluValAlaG 303
Db      3791046 C-----AATATGCTGTCTCTTCCGAGTCGTTGGCCGTCGAAAAATGCCAGC 3791093
QY      303 uGluLysGlnProTrpAlaTrp-----LeuSerCysGlyHisVal 316
Db      3791094 GAGCCATGCGGCTCTGCGTATTGAAGATCCACCCCAATCGCATCTGACCATAGGGCATCA 3791153
QY      316 LHisGlyTrpHisAsnTrpGlyHisArgSerAspThrGlnAlaAsnGluArgGluCys-- 335
Db      3791154 TCATAGTAC-----TGCGACGACATATCAGATGGCGAGTCTGTATTAATATGCGCATGG 3791207
QY      336 -----PrometCysArgThrVal 341
Db      3791208 TGCCACTTATCCGCCGACGAGAACCGTC 3791235

```

Search completed: December 9, 2004, 19:16:03
Job time : 6915 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2004, 13:11:42, Search time 769 Seconds
(without alignment)

2794.369 Million cell updates/sec

Title: US-10-041-030-4

Sequence: 1 MFSPGQEHCAPNKEPVKKG.....ATQVGEONCIRLFGQPID 420

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:
-MODE=frame_plus_p2n.model -BEV=x1p
-O=cn2.1/USFTO.spool.p/US10041030/runtat.09122004.095516.7653/app.query.fasta_1.583
-DB=N.Geneseq.23Sep04 -QWTF=fastad -SUPPL=eng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10041030 @CGN 1.1 885 @runtat.09122004.095516.7653 -NCPU=6 -ICPU=3
-NO MMAP -LARGEDUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-BEV TIMEOUT=120 -MMRN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N.Geneseq.23Sep04:*

1: geneseqn19808:*\n2: geneseqn19908:*\n3: geneseqn20008:*\n4: geneseqn20018:*\n5: geneseqn20018:*\n6: geneseqn20028:*\n7: geneseqn20028:*\n8: geneseqn20038:*\n9: geneseqn20038:*\n10: geneseqn20038:*\n11: geneseqn20038:*\n12: geneseqn20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2290	100.0	1565	6	AAL46323 Human M33
2	2290	100.0	1823	10	ADC30209 Human nov
3	2290	100.0	4563	6	ABQ78319 Nucleoid
4	2290	100.0	5579	8	ACC42349 Human MAP
5	2283	99.7	1263	6	AB158449 Human pel
6	2283	99.7	1263	12	ADP48670 Human pel

Seq ID #3
Lined

7	2175.5	95.0	1260	6	AB158448	AD158448	Murine pel
8	2175.5	95.0	1260	12	ADP48668	ADp48668	Murine pel
9	2111.5	92.2	1717	8	ACC42348	ACC42348	Mouse MAP
10	1917	83.7	1257	6	AB158447	AD158447	Human pel
11	1917	83.7	1257	12	ADP48666	ADp48666	Human pel
12	1917	83.7	1304	8	ABX05095	ABx05095	Human nov
13	1917	83.7	3222	6	AAL46320	AAL46320	Human M30
14	1917	83.7	3526	6	AAL46319	AD146319	Human M30
15	1917	83.7	7136	12	ADP483102	ADp483102	Human pel
16	1916	83.7	2736	6	AAL46316	AD146316	Rat M30 c
17	1912	83.5	1257	6	AB158446	AD158446	Murine pel
18	1912	83.5	1257	12	ADP48664	ADp48664	Murine pel
19	1822	79.6	3846	6	AAL46317	AD158452	Human M30
20	1642	71.7	1338	6	AB158452	AD158452	Human pel
21	1642	71.7	1338	12	ADP48674	ADp48674	Human pel
22	1642	71.7	2683	10	ADD71199	ADd71199	Human int
23	1640	71.6	2589	6	AAL46322	AD146322	Human M31
24	1633	71.3	2588	10	ADC30836	ADc30836	Human nov
25	1632	71.3	1936	6	AAL46321	AD146321	Murine M3
26	1618	70.7	3254	6	AAL46318	AD146318	Human M30
27	1605	70.1	3332	3	ACC66453	ACC66453	Human sec
28	1574	68.7	2866	4	AAH16312	AAH16312	Human CDN
29	1574	68.7	2866	4	AAH77808	AAh77808	Nucleotid
30	1482.5	64.7	2508	10	ADB63804	ADb63804	Human CDN
31	1370	59.8	2183	4	AAK94855	AAK94855	Human ful
32	1370	59.8	2183	12	ADL31993	ADl31993	Pull lengt
33	1302	56.9	2966	4	ABL09073	ABl09073	Drosophill
34	1291	56.4	16092	4	AAK73420	AAK73420	Human imm
35	1172	51.2	803	4	AAH06690	AAh06690	Human CDN
36	1102	48.1	592	10	ADC32115	ADc32115	Human nov
37	1075	46.9	723	5	AAK68681	AAK68681	DNA encod
38	1064	46.5	567	6	AAL46314	AD146314	Human M33
39	1050	45.9	570	6	AAL46324	AD146324	Human M32
40	1022	44.6	541	12	ACH89787	ACH89787	Human ger
41	996	43.5	31882	4	ABL09072	ABl09072	Drosophill
42	974	42.5	528	12	ACH87309	ACH87309	Human gen
43	904.5	39.5	869	10	ADC32585	ADc32585	Human nov
44	858.5	37.5	741	4	AAK93882	AAK93882	Human CDN
45	858.5	37.5	741	4	AAK92298	AAK92298	Human CDN

ALIGNMENTS

RESULT 1	AAL46323	standard; cDNA; 1565 BP.
XX	XX	XX
XX	XX	XX
AC	AAL46323;	
XX	XX	XX
DT	19-UTR-2002 (first entry)	
XX	XX	XX
DE	Human M33 coding sequence SEQ ID NO: 15.	
XX	XX	XX
KW	Neurodegenerative disease; M30; M31; M32; M33; stroke;	
KW	fragile X syndrome; Huntington's disease; Parkinson's disease;	
KW	Alzheimer's disease; multiple sclerosis; ovarian cancer; allergy;	
KW	neurodegeneration; immune disorder; autoimmune disease; infection;	
KW	leukemia; inflammation; neuroprotective; cerebroprotective;	
KW	immunosuppressive; cytotoxic; neurotropic; antiparkinsonian; antiasthmatic;	
KW	virulence; antiinflammatory; gene; ss.	
OS	Homo sapiens.	
XX	XX	XX
PN	WO200221138-A2.	
XX	XX	XX
PD	14-MAR-2002.	
XX	XX	XX
PF	07-SEP-2001; 2001WO-EP010366.	
XX	XX	XX
PR	07-SEP-2000; 2000US-00657479.	
XX	XX	XX
PA	(AXAR-) AXARON BIOSCIENCE AG.	
XX	XX	XX

XX 24-SEP-2002-2002NC09030474.
 XX 24-SEP-2001-2004US-0324634P.
 XX (HYSE-) HYSEQ INC.
 XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T,
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G,
 PI Haley-Vicente D, Drmanac RT;
 DR WPI: 2003-371981/35.
 DR P-PSDB; ADC31180.
 XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX claim 1; SEQ ID NO 291; 1185bp, English.
 XX The invention relates to 971 novel human cDNA sequences (ADC329919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostic, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridization probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 1823 BP; 425 A; 483 C; 530 G; 385 T; 0 U; 0 Other:
 SO
 Alignment Scores:
 Pred. No.: 8.02e-200 Length: 1823
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-10-041-030-4 (1-420) x ADC30209 (1-1823)
 QY 1 MetPheSerProGlyngInguLniSCyValaProAsnlygInuProVallytyrGly 20
 DB 265 AAGTTTTCCTCCCTGGCCAGGAGAACACTGCGCCCAATTAAGAGCCAGCAATATCGG 324
 QY 21 GluLeuValValLeuGlytyrAanglyValaLeuProAsnGlyAspArgGlyATGAGT 40
 DB 325 GAGCTGTGTGTCTCGGTAATGCTTTCACCAATGAGATAGAGGACCGAGGAAA 384

QY 41 SerArgPheAlaLeuTyrLeuArgProLyValaAnglyValLyPProSerThrValHis 60
 DB 385 ACTAATATTTGCTCCCTTACAAAGCGGCCCAAGGCAAAATGGGTCAAAACCCAGCACCGTCCAT 444
 QY 61 ValIleSerThrProGlnAlaSerLyValaIleSerCyLySerGlyGlnHisSerIleSer 80
 DB 445 GTGATATCCAGCCGCCAGGATCCAAAGGCTTACAGCTGCACAAAGGTCAACACAGTATATCC 504
 QY 81 TyrTrpLeuSerArgAsnGlnThrValValGlyTyrThrHisAspLyAspThrAsp 100
 DB 505 TACACTTGTCAAGGATCAGACTGTGGTGGTGAATACACACATGATAGATACGAT 564
 QY 101 MetPheGlnValGlyArgSerThrGlySerProIleAspPheValValThrAspThrIle 120
 DB 565 AAGTTTCAGTGGGAGATCAACAAGCCCTTATCGACTTGGTGTCAACACAGATT 624
 QY 121 SerGlySerGlnAspThrAspGlnValaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 625 TCTGGAGCCCAAGACCGGACGAGCCAGATCACACAGACACCATATCCAGTTCGCC 684
 QY 141 CyArgIleValCyAspArgAngInuProTyrThrAlaArgIlePheAlaIleGlyPhe 160
 DB 685 TCCAGGATCGTGTGCGACAGAAATGAACTTACAGACGCGATATTGCGCCCGGATTT 744
 QY 161 AspSerSerLyAsnIlePheLeuGlyGlyValaIleValyTTrpLyAspProAspGly 180
 DB 745 GACTCTTCCAAAACATATTTCTTGAGAAAAGGACAAAGTGAAGAAAACCCCAAGCGC 804
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValaLeuValMetHisProArgGlyGlyPheThr 200
 DB 805 CACATGATGGCTCACTCATATGAGCTCTGTGATGATCATCCAGAGGGGGCTTCACC 864
 QY 201 GluGlySerGlnProGlyValIleTrpArgGlnIleSerValCyGlyAspValIleThrLeu 220
 DB 865 GAGGAGTCCCAAGCCCGGGGTCTGGCGGAGATCTGTCTGTGGAGATGTATACACTTG 924
 QY 221 ArgGlnThrArgSerThrAlaGlnIleArgGlyLyLeuValaGlySerGlyThrAsnValLeu 240
 DB 925 CCAGAAACAGATCGGCGCCAGCAAGAGAAAGCTGTGGAAAGTGAGACCAACGCTCG 984
 QY 241 GlnAspGlySerLeuIleAspLeuCyGlyValaThrLeuThrArgThrAlaAspGly 260
 DB 985 CAGAGGGCTCTCTCATTTGACCTGTGTGGGCCACTTCTCTGGAGAACACAGCATGGG 1044
 QY 261 LeuPheHisThrProThrGlnLyHisIleGlnAlaLeuArgGlnIleAsnAlaAla 280
 DB 1045 CTTTTTCATATCCAACTCAGAAACATAGAAAGCCCTCGGACAGGATTTAACCGCGCC 1104
 QY 281 ArgProGlnCyProValaGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLyGlu 300
 DB 1105 CCGCTCAGATGCTCTGTGGGCTCAACACCTCGGCTTCCCAAGATCAACAGAAAGG 1164
 QY 301 ValIleGlnGlyLeuGlnProThrAlaTyrIleuSerCyGlyValaIleGlyTyrHis 320
 DB 1165 GTGTGAGAGAAAGACGCTGGGCATATCTCAAGTTGTGGCCAGCTGACGGATCCAC 1224
 QY 321 AsnTrpGlyHisArgSerAspThrGlnAlaAngInuArgGlyCyProMetCyAspArgThr 340
 DB 1225 AACTGGGGCCATCGGAGTACACGAGGCCCAAGAGGAGTGTCCCATGTGCGAGACT 1284
 QY 341 ValGlyProTyrValProLeuTrpLeuGlyCyGlyValaGlyPheTyrValAspIleGly 360
 DB 1285 GTGGGCCCCATATGTGCTCTGTGGCTGTGAGGACGAGATTTATGTATAACCGAGGA 1344
 QY 361 ProProThrHisAlaPheThrProCyGlyHisValCySerGlyLySerAlaLyTyr 380
 DB 1345 CCGCCAACTCATGCTTTCATCTCCCTGTGACACGCTGTCTCGGAAAGCTTCGAAATTC 1404
 QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCyAspProPheCy 400
 DB 1405 TGGTCTCAAGATCCCGTGTGCTCATGAGAACTCATTCACGCTGCTGCTTCTGT 1464

Dd	1581	CGGCAACTCAGTGGTTTCACCTCCGTGTGAGACACGCTGTGGAGAACTGCAGAAAATAC	1664
Oy	361	TTPSERGINLLEProLeuPRoHIGLYThrHisAlaPheHISAlaAlaCyseProPhCyS	400
Dd	1641	TGGTCTCGAATCCCCCTTGCCCTCATGAAGAATCATGATTTCACGCTGCCCTTTCGT	1700
Oy	401	ALAThrGlnLeuValGLYGLuGlnAsnCYsIleTyLSerLIUlePheGInGLYProILeap	420
Dd	1701	GCTACACAGCTGGTTGGGGAGACAAACTGCATCAAATTAATTTTCCAAGTCAATTGAC	1760
RESULT 4			
ID	ACC42349		
XX	ACC42349 standard; cDNA; 5579 BP.		
XX			
XX	ACC42349;		
DT	22-MAY-2003 (first entry)		
DE	Human MAP kinase cascade activator #59 cDNA.		
XX			
KW	Human, Elki phosphorylation; Elki phosphorylation kinase; virucide;		
KM	antiinflammatory; immunomodulator; cytostatic; antiallergic; anti-HIV;		
KW	antirheumatic; antiarthritis; antidiabetic; antiaesthetic; gene therapy;		
KM	inflammation; autoimmune disease; viral disease; cancer; diabetes;		
KW	rheumacoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;		
XX	Iga nephritis; gene; ss.		
XX			
OS	Homo sapiens.		
PN	MO2003008569-A1.		
PD	30-JAN-2003.		
PF	15-JUL-2002; 2002WO-JP007174.		
PR	18-JUL-2001; 2001JP-00218204.		
PR	31-AUG-2001; 2001JP-00263450.		
PR	21-JAN-2002; 2002JP-00012176.		
PA	(ASAH) ASAMI KASEI KOGYO KK.		
PI	Mateuzaki O, Mateuda A, Nagano Y, Suzuki N;		
DR	WP1; 2003-229582/22.		
DR	P-PsDB; ABR41083.		
PT	Elki phosphorylation-associated gene and its encoded protein with MAP		
PT	kinase cascade effect, applicable in diagnosis of and developing drugs		
PT	for e.g. inflammations, autoimmune diseases, viral diseases and cancer.		
PS	Claim 4; Page 631-640; 762pp; Japanese.		
XX			
XX	The invention relates to a novel purified protein having Elki		
CC	phosphorylation activity and/or an activity of activating Elki		
CC	phosphorylation kinase. A protein of the invention has antiinflammatory,		
CC	immunomodulator, virucide, cyostatic, antiallergic, antirheumatic,		
CC	antiarthritic, antidiabetic, antiaesthetic, and anti-HIV activity. The		
CC	polynucleotides may have a use in gene therapy. The gene and its encoded		
CC	protein are applicable in diagnosis of and developing drugs for e.g.		
CC	Inflammatory, autoimmune diseases, viral diseases and cancer such as		
CC	hepatitis and Iga nephritis. The present sequence is used in the		
CC	exemplification of the invention		
SQ	Sequence 5579 BP; 1535 A; 1111 C; 1214 G; 1719 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No. :	3,69e-199	Length:	5579
Score:	2290.00	Matches:	420
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	B	Gaps:	0

QY	1	MetPheSerProGluYngInGluGluHnIbCyAlaProAbnlyRgluProValIlyTyTcyL	20
Db	177	ATGTTTTCCTCCGCGCAGAGAAACACTGCCCCCCCATTAAGAGCCAGTAAATACCGG	236
QY	21	GIueuValValIleuGIyTyTAsnGIyAlaIeupProAbnGIyAbaPAsGIyArGaRgYs	40
Db	237	GAGCTGGTGGTCTCGGGTACAAATGGTCTTTACCCAATGAGATAGAGACGAGAGAA	296
QY	41	SeArGPheaIaIeUyTyTAsnGPProIyAaIaAbnGIyValIyVProSeTThrValHs	60
Db	297	AGTAGATTGCCCCCTCTACAAAGCGGCCCAAGCAATGGTGTCAAAACCAGCACCGTCCAT	356
QY	61	ValIleSerThrProGluAlaIleSerIyAlaIleSerCyblyRglYngInHnIleSerIleSer	80
Db	357	GTGATATCCAGCCCCCAGCGCATCCAGAGCTATCAGCTGCAAAAGGTCAACAGATATACC	416
QY	81	TyTrhIleuSeArGaSnGIuThValValValGIuTyTrhHnIbAsPlyAsPThArP	100
Db	417	TACACTTTGTCAAGGAATCAGACTGTGTGTGGAGTACACACTGATAGATACGAT	476
QY	101	MetPheGIuValGIyArSeTThrGIuSeRProIleAaPheValValThnAsPThrIle	120
Db	477	ATGTTTCAAGTGGGAGATCAACAGAAAGCCCTATCGACTTCGTTGTCAACAGACATY	536
QY	121	SeRglYSeRGIuAenThrAsPGIuAlaGlnIleThrGlnSerThrIleSeArGPheIa	140
Db	537	TCTGGCACCCAGAAACACGAGAAAGCCAGATCACACAGACACCATATCCAGTTCCGC	596
QY	141	CysArGIleValCyAsAPArAsnGIuProTyTrhAlaArGIlePheAlaIaGIyPhe	160
Db	557	TGCAGATCTGTGTGCGACAGAAATGAACCTTACACAGACCGATTTGCGCCGCGATT	656
QY	161	AaPSeSerIyAaAnIlePheIeUGIyGIuIyAaIaIyTyTrPlyAaSnProAaPGIy	180
Db	657	GACTCTTCCAATAACATATTTCTTGAGAAAGACAGCAAGGTGAAATAAACCCCGACGGC	716
QY	181	HnIleuArPGIyLeuThrThrAsnGIyValIleuValIeThnIbAsPProArGIyGIyPheThr	200
Db	717	CACATGATGGGCTCACACTAAATGCGCTCTGTGTGAAGCATCCACGAGGGGCTTCAAC	776
QY	201	GIuGIuSeRGIuProGIyValTyTrPArGIuIleSeRValCySbGIyAsPValTyTrhIeU	220
Db	777	GAGGAGTCCAGCCCGGGGTCTGGCGCGAGATCTCTGTGTGAGATGTGTACACTTG	836
QY	221	ArGIuThrArGSerIaIeGIuIaRGIyLyIeAaValGIuSeRGIuThrAsnValIeU	240
Db	837	CGAGAAACAGTGGCGCCAGCAACAGAGAAAGCTGGCGAAAGTGAACCAACGTCCTG	896
QY	241	GIuAaPGIySerIleuIleAaPLeuCyGIyAlaThrIleuLeuThrPArGIuThrIaAaPGIy	260
Db	897	CAGACGGCTCTCCATYTAGCTGTGTGGGGCCACTCTCTCTGAGAAACGCAAGATGG	956
QY	261	LeuPheHnIbThrProThrGIuIyHnIleGIuAlaIeUArGIuGIuIleAaAlaIa	280
Db	957	CTTTTCTACTCCAACTCAGAAGCACTATAAGCCCTCCGCGAGAGATTACGCCCC	1016
QY	281	ArGPProGIuCyProValGIyLeuAaSnThrIleuAlaPheProSerIleAaArGIyGIu	300
Db	1017	CGGCTCAGTGTCTGTGGGGCTCAACACCTGGCCTTCCCAACATCAACAGAGAAAG	1076
QY	301	ValValGIuGIuIyGIuProTPrAlaTyTrIeUSeRGIyGIuIleValHnIleGIyTyTrHs	320
Db	1077	GTGGTGGAGAGAAAGACCCCTGGCGCATATCTCAATTGTGGCCACGATGACAGGGATACAC	1136
QY	321	AaenTyRGIyHnIbArGSerAsPThrGIuAlaAbnGIuArGIuCyProMetCyArGIuThr	340
Db	1137	AACGTGGGCCATCGAGATGACACGAGGCCCAACGAGAGGAGTGTCCCATGTGCAGACT	1196
QY	341	ValGIyProTyValProLeuTyRleuGIyCyGIuAlaGIyPheTyValAaPAlaGIy	360

Dd	1197	GTGGGCCCCCTATGTCGCCCTCTCGCTTGCGCTTGTAGGACGAGATTTTATGTATGACGCACGA	1256
Oy	361	ProDrothrhIsalapherThrProCysglYhiSvalCySeSerGIuIySeSerAlalyTyR	380
Dd	1257	CCGGCAACTCATGCTTTTCACCTCCCTGTGTGACACCGTGTCTCGGAAAGTCTGCAAAATVAC	1316
Oy	381	TyrSerGniIleProlLeuProHISgLYlhrhIsAlapHehIsAlaICyAProPhoCys	400
Dd	1317	TGTGTTCAGATCCGCTTCCCTCATAAGAATCATGATGATTTCACGCTGCTTGCCTTTCTGT	1376
Oy	401	AlAthrgInLeuValIGlyGLUGlnAsmCYsilleySleuilepheGnIGlyProileasp	420
Dd	1377	GCTACACAGCTGGTGTGGGACCAAACATGCATCAATTAATTTTCCAAGTCCAAATTGAC	1436
RESULT 5			
ID	ABL58449	standard; DNA; 1263 BP.	
XX	ABLS8449;		
XX	30-JUL-2002	(first entry)	
DE	Human pellino-2 polypeptide coding sequence.		
XX			
KW	Nuclear factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; virucide;		
KM	antibacterial; fungicide; protozoocide; antischismatic; antihematic;		
KW	antiarthritic; antiinflammatory; antiatherosclerotic; neuroprotective;		
KM	nocrotropic; anticulcer; human; pellino-2; gene; ds.		
XX			
OS	Homo sapiens.		
FH			
FT	CDS	Location/Qualifiers	
FT		1..1263	
FT		/+tag= "a"	
FT		/product= "pellino-2"	
XX			
EN	WO200183739-A2.		
XX			
PD	08-NOV-2001.		
XX			
PF	27-APR-2001; 2001WO-US013676.		
XX			
PR	28-APR-2000; 2000US-0200198P.		
XX			
PA	(IMMV) IMMUNEX CORP.		
PI	Bird TA, Cosman DJ:		
DR	WPI; 2002-066532/09.		
XX	P-PSDB; ABB07922.		
PT			
PT	New Pellino polypeptides for identifying compounds that alter polypeptide		
PT	activity, treating pathogenic infection or inhibiting apoptosis, are		
PT	capable of stimulating nuclear factor-kappaB- or p38-dependent		
PT	transcription.		
XX			
PS	Claim 3; Page 61; 70pp; English.		
XX			
CC	The invention provides polypeptides capable of stimulating nuclear factor		
CC	(NF)-kappaB-dependent transcription or p38-dependent transcription,		
CC	referred as Pellino polypeptides. The pellino polypeptides are useful for		
CC	identifying modulators that alter the pellino polypeptide and pellino		
CC	dominant-negative activity. They are also useful for identifying		
CC	compounds that inhibit the binding activity of the polypeptides and to		
CC	studying cell-signal transduction. They are useful for preventing or		
CC	treating infection by a pathogen such as virus, bacterial, fungi, algae		
CC	or protozoa, or inhibiting apoptosis. Dominant-negative pellino		
CC	polypeptides are useful for treating inflammatory conditions such as		
CC	asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's		
CC	disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and		
CC	also for inhibiting mitogen activated protein (MAP) kinase-activated		
CC	pathways. Pellino polypeptides and polymunocleotides are useful to identify		
CC	small molecule inhibitors of protein association or function of Pellino,		

Sequence 1260 BP; 309 A; 357 C; 359 G; 235 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Mismatches:	Indels:	Gaps:
1.52e-189	1260	2175.50	399	10	1
Score:	2175.50	Conservative:	10		
Percent Similarity:	97.38%	Mismatches:	10		
Best Local Similarity:	95.00%	Indels:	1		
Query Match:	95.00%	Gaps:	1		

US-10-041-030-4 (1-420) x ADP48668 (1-1260)

```

QY 1 MetPheSerProGlyGlnGluGlnHisCySAIaProAAsnLysGluProValIleYrGly 20
Db 1 ATGTTTTCCTCCGGCCAGAGAGAACCCACCCCAACAAGAGCGGTGAATACGGG 60
QY 21 GLeuValValIleuGlyTyrAAsnGlyAlaLeuProAAsnGlyAAspArgIleYr 40
Db 61 GAGCTGGTGTCTGGGGGTACAAATGCTTAACTTAATGTGACAGGGGCGAGGAAA 120
QY 41 SerArgPheAlaLeuTyrIleYrSApProYSAIaAsnGlyValIleProSerThrValHis 60
Db 121 AGCAAGATTGCGCTTAATAGCGGACCTACGCCAGGTGTGTAACCAACCCAGCAATCCAC 180
QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCySAIeGlyGlnHisSerIleSer 80
Db 181 ATGCTTCACACACAGAGCGCTCCAGGCCATCAGCTCAGAGAGACATACACATATCG 240
QY 81 TyrThrLeuSerArgAAsnGlnThrValValIleGluTyrThrHisAspIleYrAspThrAsp 100
Db 241 TACACGTTGTACAGGAGCAGACAGCGTAGGTGAGTACACACATATTAAGACACGAGAC 300
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValIleThrAspThrIle 120
Db 301 ATGTTTCAGGTGGCGGTCAACAGAAACCCCATTAAGTCTGTGTCAACAGACAGGT 360
QY 121 SerGlySerGlnAAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 361 TCCGGCGGTCAAGACGAAAT---GCCAGATCACACAGACCACTCTAGTTGGTTCGA 417
QY 141 CySAIleValCySApArgAAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 418 TCCAGGATCGTGTGACAGAAACGAGCATATACAGACGATATTCGCGGACGATTC 477
QY 161 AspSerSerLysAAsnIlePheLeuGlyGluValAlaIleTyrTrpLysAsnProAspGly 180
Db 478 GATTCTTCAAAAATATCTTTCTTGAGAGAAAGCAAAAATGAAAAACCTGTATGGA 537
QY 181 HisMetAspGlyLeuThrThrAAsnGlyValIleValMetHisProArgGlyGlyPheThr 200
Db 538 CACATGAGTGAAGCTCACTCAATGCTGTCTAGTATGACACCCGCAAGAGAGCTTCAC 597
QY 201 GLeuLysSerGlnProGlyValIleTyrArgGluIleSerValCySAIeYrAspValIleYrThrLeu 220
Db 598 GAGGATCCAGGCTGAGCTGAGAGAGATCTCTGTCTGTGGGATGTGACACCTTG 657
QY 221 ArgGluThrArgSerAlaGlnGlnIleArgLysValIleValGluSerGluThrAsnValIleu 240
Db 658 CGAGAACCAAGGTCGCGCCAGAGAGAGAGAGTGTGTGAAGTGAAGCAACGCTCG 717
QY 241 GlnAspGlySerLeuIleAspLeuCySAIeAlaThrLeuThrTyrArgThrAlaAspGly 260
Db 718 CAAGACGGCTCCTCATTAACCTGTGTGGGCACTCTCTCTGAGGAACCCCAAGATGGC 777
QY 261 LeuPheHisThrProThrGlnLysHisIleGlnAlaLeuArgGlnGluIleAsnAlaAla 280
Db 778 CTTTTTCACGCTCTCTACAGAGCAATAGAAAGCCCTCCGCGAGAGATCATGACGCC 837
QY 281 ArgProGlnCySApProAlaGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
Db 838 CGAGCCCAAGTGGCCCTGTGGCTTAAACCCCTGTCCCTCCAGCATCAACCGGAAGGA 897
QY 301 ValValGluGluGlnProThrAlaTyrLeuSerCySAIeHisValHisGlyTyrHis 320

```

```

Db 898 GTGCTGAGAGAGAGACACCTGGGCATACCTGAGCTGGGCGCATGTCCAGGGCTACAC 957
QY 321 AAsnTyrGlyHisArgSerAspThrGlnAlaAsnGluTyrGluCySApMetCySApThr 340
Db 958 AGCTGGGGCCATCGGAGCGAGCGGAAACCAACGAGAGAGATGTCCCATGTGCGAGACT 1017
QY 341 ValGlyProTyrValProLeuThrLeuGlyCySAIeGlnAlaGlyPheThrValAspAlaGly 360
Db 1018 GTGGGCCCCCTACGTCCTCTCTGTGGCTGTGAGGAGAGATTTTATGTGATCGGGA 1077
QY 361 ProProThrHisAlaPheThrProCySAIeHisValCySApSerGluYrSerAlaIleYr 380
Db 1078 CCCCCAATCAACGCTTTCACCCCTGCGGACCGCTGTTCAGAAATCTGCGCAATGAC 1137
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCySApProPheCys 400
Db 1138 TGTGTGCAAGATCCCATGCCCCCAACGAGCGGTTTCATGTGCGCTGTCTCTGTGC 1197
QY 401 AlaThrGlnLeuValGlyGluGlnAAsnCySAIeLysLeuIlePheGlnGlyProIleAsp 420
Db 1198 GCCACGACGCTGTGTGTGTAACAGACTGCATCAATGATTTTCCAAAGTCCAGTGCAC 1257

RESULT 9
ACC42348
ID ACC42348 standard; cDNA; 1717 BP.
XX
AC ACC42348;
XX
DT 22-MAY-2003 (first entry)
XX
DE Mouse MAP kinase cascade activator #9 cDNA.
XX
KW Mouse, Etk1 phosphorylation; Etk1 phosphorylation kinase; virucide;
KW antiinflammatory; immunomodulator; cytolethal; antiallergic; anti-HIV;
KW antirheumatic; antidiabetic; antidiabetic; antidiabetic; gene therapy;
KW inflammation; autoimmune disease; viral disease; cancer; diabetes;
KW rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
KW IGA nephritis; gene; ss.
XX
OS Mus musculus.
XX
PN W02003008589-A1.
XX
PD 30-JAN-2003.
XX
PF 15-JUL-2002; 2002MO-JP007174.
XX
PR 18-JUL-2001; 2001JP-00218204.
PR 31-AUG-2001; 2001JP-00263450.
PR 21-JAN-2002; 2002JP-00012176.
XX
PA (ASAH) ASAH KASEI KOGYO KK.
PI Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;
XX
DR MPI: 2003-229582/22.
DR P-PSDB; ABR41082.
XX
PT Etk1 phosphorylation-associated gene and its encoded protein with MAP
PT kinase cascade effect, applicable in diagnosis of and developing drugs
PT for e.g. inflammations, autoimmune diseases, viral diseases and cancer.
XX
PS Claim 4, Page 623-627; 762pp; Japanese.
XX
CC The invention relates to a novel purified protein having Etk1
CC phosphorylation activity and/or an activity of activating Etk1
CC phosphorylation kinase. A protein of the invention has antiinflammatory,
CC immunomodulator, virucide, cytolethal, antiallergic, antirheumatic,
CC antidiabetic, antidiabetic, antidiabetic, and anti-HIV activity. The
CC polynucleotides may have a use in gene therapy. The gene and its encoded
CC protein are applicable in diagnosis of and developing drugs for e.g.
CC inflammations, autoimmune diseases, viral diseases and cancer such as

```


XX 11-DEC-2002; 2002US-00317250.
 PR (IMV) IMMUNEX CORP.
 PA (CLEV-) CLEVELAND CLINIC FOUND.
 XX
 PI Bird TA, Cosman DJ, Li X;
 DR WPI; 2004-480927/45.
 DR P-PSDB; ADP48667.
 XX
 PT Identifying inhibitors of IL-1 signaling, useful for treating e.g.,
 PT asthma, and rheumatoid arthritis, comprises assaying the association of a
 PT pellino-1 polypeptide with IRAK-4 in the presence of a test compound.
 XX
 PS Claim 9; SEQ ID NO 3; 81pp; English.
 XX
 CC The present invention describes a method for identifying compounds (C)
 CC that inhibit nuclear factor kappa B (NF-kB)-dependent transcription or
 CC p38-dependent transcription. The method comprises: (1) mixing a test
 CC compound with a Pellino-1 polypeptide; (b) assaying the association of
 CC the Pellino-1 polypeptide with IL-1 receptor-associated kinase 4 (IRAK-4)
 CC in the presence of the test compound; and (c) determining whether the
 CC test compound inhibits the association of the Pellino-1 polypeptide with
 CC a binding partner. Also described: (1) an inhibitory nucleic acid that
 CC binds to a nucleic acid encoding an amino acid sequence as described
 CC above, where the presence of the inhibitory nucleic acid within a cell
 CC inhibits the association of Pellino-1 with IRAK-4; and (2) an inhibitory
 CC polypeptide comprising an antibody fragment that binds to a polypeptide
 CC comprising an amino acid sequence as described above, where the presence
 CC of the inhibitory polypeptide within a cell inhibits the association of
 CC Pellino-1 with IRAK-4. (C) have anti-rheumatoid, anti-arthritis,
 CC antiasthmatic, antiinflammatory, antirheumatic, antitumor,
 CC gastrointestinal, neuroprotective and nociceptive activities, and can be
 CC used as an IL-1 antagonist, IRAK protein kinase family inhibitor, and NF
 CC kappa B inducing kinase inhibitor. The methods and compositions of the
 CC present invention are useful for the prevention and/or treatment of
 CC diseases or conditions associated with aberrant expression or activity of
 CC the Pellino-1 polypeptide, such as asthma, rheumatoid arthritis,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC atherosclerosis and Alzheimer's disease. The present sequence encodes
 CC human Pellino-1, which is used in the exemplification of the present
 CC invention. The human Pellino-1 gene is located on chromosome 2, more
 CC specifically to 2p13.3.
 CC
 XX
 SQ Sequence 1257 BP; 368 A; 264 C; 296 G; 329 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 7,49e-166 Length: 1257
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.48% Conservative: 38
 Best Local Similarity: 81.43% Mismatches: 38
 Query Match: 83.71% Indels: 2
 DB: 12 Gaps: 1
 US-10-041-030-4 (1-420) x ADP48666 (1-1257)
 QY 1 MetPheSerProGlyInGluGluIuHisCyAlaProAbnlySGluProValIlyrTYGly 20
 DB 1 ATGTTTTCCTCATCAAGAAATCAT-----CAATCTTAAAGCACACAGTAAATTTGGT 54
 QY 21 GluLeuValIleuGlyTYrAsnGlyAlaLeuProAbnGlyAspArgGlyArgGly 40
 DB 55 GAACCTATGCTCTTAGATTAATGATCTCTCCAAACCGTATGAGAGAGAGAGAA 114
 QY 41 SerArgPheIleuTyrlsAspArgProValAsnGlyValIlyrProSerThrValHis 60
 DB 115 AGTAGGTTTGCTTGTATTAAGACCTTAAGGCAAAAGGGGTGAAGCCACAGCTGTCAT 174
 QY 61 ValIleSerThrProGlnAlaSerIlyValIleSerCylsGlyGlnIHisSerIleSer 80
 DB 175 ATGCTTGTACTCTCTCAAGGCTGCAAGGCAATTAAGCAACAAACACAGATTAATCA 234

QY 81 TyrThrLeuSerArgAsnGlnThrValIleValGluTyThrHisAspIlyAspThrAsp 100
 DB 235 TATACCTTTATCTCGGGCCACAGCTGTGTGTGAATATATCATGACAGCAACACCGAT 294
 QY 101 MetPheGlnValIlyrAspSerThrGluSerProIleAspPheValIlyrAspThrIle 120
 DB 295 ATGTTTCAGATTGGCCGGTGCAGCTGAAGCCCATTTGATTTGTATTAAGTACAGCGTT 354
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 355 CCGAAGTCAAGTCAAGTATTCGTATACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 414
 QY 141 CyAlaGlyIleValIyAspArgAsnGluProTyThrAlaArgIlePheAlaAlaGlyPhe 160
 DB 415 TGCAGATATATGTGAAGGAAATCTCTTTCAGCAGCAAGGATTTATCTCGAGGCTTT 474
 QY 161 AspSerSerIlyAsnIlePheLeuGlyIuValAlaIlyrTrpIlyAsnProAspGly 180
 DB 475 GACTCATCAAAAACATCTTTCTTGGGAGAAAGCTGCCAAATGGAAACATCAGATGGA 534
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleuValMetHisAspArgIlyGlyPheThr 200
 DB 535 CAGATGATGGCTTGACCACTAATGATGTTCTTTGTATGATCATCCAGCAATGGTTCA 594
 QY 201 GluIlySerGlnProGlyValIlyrArgIlyIleSerValIyGlyAspValIlyrThrLeu 220
 DB 595 GAAGACTCCAGGCTGGAATATGAGAAATATCGTCTGTGGAAATGATATTACCTTA 654
 QY 221 ArgGluThrArgSerAlaGlnGlnIlyrGlyIlyValIySerGlnIlyrAsnValIleu 240
 DB 655 GTGAAACCAAGATCGCTCAAGCAAGAAAGAAATGTGAAATTAACCAATCACTGTTA 714
 QY 241 GlnAspGlySerLeuIleAspLeuCyGlyAlaThrLeuLeuTrpArgThrAlaAspGly 260
 DB 715 CAAGATGGCTCGTTATTAATGACCTGTGTGCAACTGTATTAGGCTCTGCAAGAGG 774
 QY 261 LeuPheHisThrProThrGlnIlyHisIleGluAlaLeuArgGlnIlyLeuAlaIle 280
 DB 775 CTTTCCACACTCTTCACTGGAAGCTTTAGAGCTTTAAGACAGAAATCAATGACAGA 834
 QY 281 ArgProGlnCyProValIyGlyLeuAsnThrAlaIlePheProSerIleAsnArgIlyGlu 300
 DB 835 CAGCTTCAAGTCCCTTAGGCTTCAACACACTGACATTTCTTGTATGAAGAGAAAGAC 864
 QY 301 ValIleGluGluIySerProTrpAlaTyrlsSerCyGlyHisIyValIleGlyTyThrHis 320
 DB 895 GTTGTAGATGAAGAAACACATGGGTATATCTTAACCTGGCGCATGATGCTATCAT 954
 QY 321 AsnTrpGlyHisIySerAspThrGluAlaAsnGluArgGluCyProMetCyAlaGlyThr 340
 DB 955 AACTGGGGAAACAAAGAAAGAGTGTGCAAAAGATCGGAATGTCCTTATGTGTAGTCT 1014
 QY 341 ValIlyProTyValIyProLeuTrpLeuGlyIyGluAlaGlyPheTyThrValIleAspIly 360
 DB 1015 GTTGTGCTTATGTTCTCTGTGGCTTGAATGAACTGAGATTTTATGTGAGCGCCGC 1074
 QY 361 ProProThrHisAlaPheThrProCyGlyHisIyValIySerGlnIlySerAlaIlyr 380
 DB 1075 CCGCAACCCAGATCGTTAGCCCGTGGGCAATGTGTGTGAAAGAAACAACTGCTAT 1134
 QY 381 TrpSerGlnIleProLeuProHisIyGlyThrHisAlaPheHisAlaIlyCyProPheCy 400
 DB 1135 TGGTCCCAAGTCCCACTTCTCATGATGATCTCACTTTCACAGCAAGCCGTGCTTTTGT 1194
 QY 401 AlaThrGlnLeuValIyGluGlnAsnCyIleIyIleValIlePheGlnIlyProIleAsp 420
 DB 1195 GCAATCAATGTGGCTGTGTAACAGGCTTCAATCAAGCTTATTTTCAAGAACTCTAGAC 1254
 RESULT 12
 ID ABX05095
 AC ABX05095 standard; cDNA; 1304 BP.
 XX
 AC ABX05095;

XX 17-JAN-2003 (first entry)
 DT Human novel polynucleotide #110.
 XX
 DE Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer;
 XX neurodegenerative disorder; lymphoid cell disorder; osteoporosis
 KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
 KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;
 KW fungal infection; bacterial infection; autoimmune disease; diabetes;
 KW atopic dermatitis.
 XX
 OS Homo sapiens.
 XX
 PN WO200274961-A1.
 XX
 PD 26-SEP-2002.
 XX
 PF 14-MAR-2002; 2002WO-US005109.
 XX
 PR 15-MAR-2001; 2001US-00810173.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F,
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
 PI Mehman T, Wang J, Wang D, Drmanac RT;
 XX
 DR WPI; 2003-040556/03.
 DR P-PSDB; ABU00017.
 XX
 PT New isolated polypeptides and polynucleotides, useful for preventing,
 PT treating or ameliorating medical conditions such as cancer,
 PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
 PT disorders, and infections.
 XX
 PS Claim 1; SEQ ID NO 110; 235pp; English.
 XX
 CC The invention relates to human polynucleotides and the polypeptides they
 CC encode. The polynucleotides and polypeptides are useful in diagnostic-
 CC forensics, gene mapping, medical imaging, identification of mutations,
 CC responsible for genetic disorders or other traits, assessing biodiversity
 CC and producing many other types of data and products dependent on DNA and
 CC amino acid sequences. They are also useful for preventing, treating or
 CC ameliorating medical conditions, such as cancer, neurodegenerative
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
 CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
 CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
 CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
 CC Sequences ABX04986-ABX05511 represent human polynucleotides of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification but is based on sequence information supplied
 CC by the European Patent Office
 XX
 SQ Sequence 1304 BP; 385 A; 270 C; 305 G; 344 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 7 87e-166 Length: 1304
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.48% Conservative: 38
 Best Local Similarity: 81.43% Mismatches: 38
 Query Match: 83.71% Indels: 2
 DB: 8 Gaps: 1
 US-10-041-030-4 (1-420) X ABX05095 (1-1304)
 QY 1 MetPheSerProGlyGlnGluGluHisCySAIaProAsnLysGluProValIlySTyGly 20
 DB 1 AGTTTTCCTCGATGATCAAGAAATATCAT-----CGATCTAAAGCAACATTAATATGAT 54
 QY 21 GtLueuValIalIeugLlyTyraGngIyAlaLeuProAsnGlyAspArgGlyIargArgIys 40
 DB 55 GAACATATGTCTTAGGTAATATAGGATGCTCTCCCAATGCGCATAGAGAAAGAGGAAA 114

QY 41 SerArgPheAlaLeuTyrlYarGargProLysAlaAsnGlyValIlyProSerThrValHis 60
 DB 115 AGTAGGTTTGCTTTTTPAAGACCTAAGGCAATGGGGTGAACCCAGCATGTGAT 174
 QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCySLyGlyGlnHisSerIleSer 80
 DB 175 ATGTCTTGATCCCTCAGGCTGCAAGGCAATTAACCAAAAGACAGCATATGATATCA 234
 QY 81 TyrlrLueSerArgAsnGlnThrValValIalGluTyrlrThrlAspLysAspThrAsp 100
 DB 235 TATACCTTGTCTCGGGCCAGCATGTGTGTGAATATACATGACAGCAACAGAT 294
 QY 101 MetPheGlnValAlaYArgSerThrGluSerProIleAspPheValThrAspThrIle 120
 DB 295 ATGTTTCAGATGGCGGCTGACCTAAGCCCATTAATTTTGATGATACACGGTT 354
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrgInserThrIleSerArgPheAla 140
 DB 355 CCTGGAAGTCAAGTAATTCGATACACAGTCAGTAACAAGCATATATCAAGATTGGC 414
 QY 141 CybArgIleValCybAspArgAsnGluProTyrlrThrlAlaGlyIlePheAlaIalGlyPhe 160
 DB 415 TGCAGAAATCATATGGAACGGATCTCCCTTTACAGACCGATTTATGCTGACAGATT 474
 QY 161 AspSerSerLysAsnIlePheLueGlyGluLysAlaIalAlYTrPlyAsnProAspGly 180
 DB 475 GACTCATCAAAAACATCTTTCTTGGAGAAAGCTGCCAATTAAGACATCAGATGCA 534
 QY 181 HisMetAspGlyLueThrThrAsnGlyValIeueValMetHisProArgGlyGlyPheThr 200
 DB 535 CAGATGATGGCTGATACCATATATGTTCTTGATGATGATCCAGCAATGGGTTCA 594
 QY 201 GtLueGlnProGlyValItrPargIuIleSerValCybGlyAspValTyrlrLue 220
 DB 595 GAAGACTCAAGCTCGAATATGAGAAATATCGGTGTGGAATATTTAGCCCTA 654
 QY 221 ArgGluThrArgSerAlaGlnGlnArgLysLysLueValGluSerGluThrAsnValIleu 240
 DB 655 COTGAACCAATGCGCTCAGCAGAGAGAAAGATGCGAAATGAAACCAATCAGTTA 714
 QY 241 GlnAspGlySerLueIleAspLueCyGlyAlaThrLueLueTrpArgThrAlaAspGly 260
 DB 715 CAAGATGCTGCTTATATGACCTCTGTGTGCAATGTATGCGATCTCAAGAGGC 774
 QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLysuArgGlnIleAsnAla 280
 DB 775 CTTTCCACACCTCTAACCGTGAAGCATTTAAGACAGAAATCAATGCAGCA 834
 QY 281 ArgProGlnCybProValGlyLueuAsnThrLueAlaPheProSerIleAsnArgLysGlu 300
 DB 835 CGACCTCAGTCCCTGTAAGGTTCAACACATGATTTCTTAAGTATGAAGAGAAAGAC 894
 QY 301 ValValGluGluLysGlnProTrpAlaTyrlrLueSerCyGlyHisValHisGlyTyrlHis 320
 DB 895 GTTGTAGATGAAGAAACAACCATGGGTATATCTAACTCGCGCATGATGATGATCAT 954
 QY 321 AsnTrpGlyHisLysSerAspThrGluAlaAsnGluAlaGluCybProMetCybArgThr 340
 DB 955 AACTGGGGAACCAAGAAAGACGTATGAAAGATGCTGAATGCTCTATGTGTAAGTCT 1014
 QY 341 ValGlyProTyrlValProLueTrpLueGlyCybGluAlaGlyPheTyrlValAspAlaGly 360
 DB 1015 GTTGTCCCTATGTTCTCTGTGCTTGAATGTGAAGCTGATTTTATGTGACCCGGC 1074
 QY 361 PropThrHisAlaPheThrProCybGlyHisValCybSerGlyLysSerAlaLysTyrl 380
 DB 1075 CCTTCAACCAATGCGTTAGCCGTTGGGATGATGTGTCAAGAAAGACAACTCCAT 1134
 QY 381 TrpSerGlnIleProLueProHisGlyTyrlHisAlaPheHisAlaAlaCybProPheCyb 400
 DB 1135 TGTGTCCAGATCCCACTTCCATAGTACTCATACTTATGACAGCCGTGCTCCCTTTGT 1194

Db 1178 CCTCCAAACCCATGCGTTAGCCCGTGTGGCATGTGTGTTTCAGAAAAGCAACTGCCAT 1237
 QY 381 TTPSerginIIleProleuProHISGlyTHrHlaAlaPhenilaAlaCySerProPheCys 400
 Db 1238 TGGTCCCAATCCCACTTCCTCATGTGACTCATCTTCTCAAGCAGCCTGCTCTTTTGT 1297
 QY 401 AlaThrgIleValGlyGluGlnaenCySilelyLeuIlePheGlnGlyProIleAsp 420
 Db 1298 GCACATCACTTGGCTGTGTGAACAAGGCTACATCACTTATTTTTCAGAGACCTTACAC 1357

RESULT 14
 AAL46319
 ID AAL46319 standard; cDNA; 3526 BP.
 XX AAL46319;
 XX
 DT 19-JUL-2002 (first entry)
 XX
 DE Human M30 variant C coding sequence SEQ ID NO: 7.
 XX
 KM Neurodegenerative disease; M30; M31; M32; M33; stroke;
 KM fragile X syndrome; Huntington's disease; Parkinson's disease;
 KM Alzheimer's disease; multiple sclerosis; ovarian cancer;
 KM neurodegeneration; immune disorder; autoimmune disease; allergy;
 KM infection; leukaemia; inflammation; neuroprotective; cerebroprotective;
 KM immunosuppressive; cytostatic; nootropic; antiparkinsonian; anti-allergic;
 KM virucide; anti-inflammatory; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200221138-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 07-SEP-2001; 2001WO-EP010366.
 XX
 PR 07-SEP-2000; 2000US-00657479.
 XX
 PA (AXAR-) AXARON BIOSCIENCE AG.
 XX
 PI Schneider A, Hiemisch H, Rosner M, Klugmann M, Naim J,
 PI Eisenhardt G, Kuner R, Lananah A, Worley P, Spielvogel D, Scheek S;
 DR WPI; 2002-292287/33.
 DR P-PSDB; AAO17518.
 PT
 PT Diagnosis of neurodegenerative disease comprises detecting level of M30-
 family proteins.
 PS
 PS Claim 3; Page 91-95; 130pp; German.
 CC The present invention relates to a method of diagnosing neurodegenerative
 CC diseases, comprising determining the concentration of a protein in a body
 CC sample, where the protein may be M30 or a variant thereof, M31, M32 or
 CC M33. The method is used to diagnose neurodegenerative diseases,
 CC particularly stroke but also e.g. fragile X syndrome, Huntington's,
 CC Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also
 CC overexpression of M31 can be used for diagnosis of carcinoma and sarcoma,
 CC especially ovarian cancer. The proteins can be used to identify specific
 CC ligands, potentially useful for treating neurodegeneration, immune-system
 CC disorders (e.g. autoimmune diseases, allergy, viral infection, leukaemia,
 CC inflammation etc.), carcinoma and sarcoma. Inhibitors of the interaction
 CC between the proteins and the protein kinase IRAK-1 can be used to treat
 CC neurodegeneration. The present sequence is a coding sequence of a protein
 CC used in the method of the invention

Alignment Scores:
 Pred. No.: 3 06e-165 Length: 3526
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.48% Conservative: 38
 Best Local Similarity: 81.43% Mismatches: 38

M31, M32, M33

Query Match: 83.71% Indels: 2
 DB: 6 Gaps: 1
 US-10-041-030-4 (1-420) x AAL46319 (1-3526)

QY 1 MetPheSerProGlyGlnGluGlnHISCySAlaProaenlyGluProValIleSryGly 20
 Db 408 ATGTTTCTCCGATCAAGAAATATCAT-----CCATCTAAAGCAGGTAATAATATGCT 461

QY 21 GluLeuValIleValGlyGlyThrAsnGlyAlaLeuProaenlyGlyAspArgGlyAspArgGly 40
 Db 462 GAACTCATGTGCTTGGGTATATATGCTCTCCCAATGGCGATGAGAAAGGAGAA 521

QY 41 SerArgPheAlaLeuTyIleAspProValIleAsnGlyValIlePheProSerThrValIle 60
 Db 522 AGTAGGTTTGCTTGTATTAAGACCTTAAGGCAATGGAGGTGAGACCCAGCATGTGCT 581

QY 61 ValIleSerThrProGlnAlaSerIleValIleSerCySlyGlyGlnHISerIleSer 80
 Db 582 ATGCTTGTACTCCACAGCTGCAAGGCAATTAAGCAAAAGACACATATGATATCA 641

QY 81 TyThrIleSerArgAsnGlnThrValIleValIleGlyTyThrHISAspIleAspThrAsp 100
 Db 642 TATACTTGTGTCGGGCCAGACTGTGTGTAATACTCATGACAGCAACACAGAT 701

QY 101 MetPheGlnValIleGlySerThrGluSerProIleAspPheValIleThrAspThrIle 120
 Db 702 ATGTTTCAGATTGGCCGCTGACCTAAGCCCATATGTTTGTATGTAATGACACGGTT 761

QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 762 CCTGAAAGTCAAGATTAATTCATACACAGTCAGTACAAAGCATATATCAAGATTGCC 821

QY 141 CySArgIleValCySAspArgAsnGluProTyThrHISAlaGlyIlePheAlaIleIlePhe 160
 Db 822 TGCAGAAATCATATGGAAGGAAATCTCTTACAGACGATTAATGATGACGAGATT 881

QY 161 AspSerSerIleAsnIlePheLeuGlyGluTySAlaIleAlaIleTyIleAspProAspGly 180
 Db 882 GACTCATCAAAATCATCTTCTTGGGAGAAAGCTGCCAATGGAACATCATCATGATGA 941

QY 181 HISAspArgIleuThrThrAsnGlyValLeuValMetHISProArgGlyGlyPheThr 200
 Db 942 CAGATGATGGCTTACCATATATGCTGTTCTTGATGATGATCCACGCAATGGGTTCCACA 1001

QY 201 GluGluSerGlnProGlyValITPArgGluIleSerValCySArgIleAspValTyThrIleu 220
 Db 1002 GAAGACTCCAGCTCGAATATGAGAGAAATATCGGTGTGAGAAATGATTTAGCCCTA 1061

QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyIleuValIleGluSerGluThrAsnValIleu 240
 Db 1062 COTGAATCCAGATCGGCTCAGCAGAGAGAGAAATGCGTGAATGAAATCAATCACTTA 1121

QY 241 GlnAspGlySerLeuIleAspIleuCySArgIleValIleuLeuThrArgThrAlaAspGly 260
 Db 1122 CAAGATGCTCGTTATATGACCTCTGTGTCAACATGTTATGGCGTATGCGAGAAAGGC 1181

QY 261 LeuPheHISThrProThrGlnIleHISIleGluAlaLeuArgGlnGluIleAsnIleAla 280
 Db 1182 CTTTCCACACTCCCTACCGTGAAGCATTTAAGCTTTAAGACAGAGAAATCAATGACACA 1241

QY 281 ArgProGlnCySProValIleuAsnThrIleuAlaPheProSerIleAsnArgGlySylu 300
 Db 1242 CGACTCATGTGCTCCGTAGGCTTCAACACATGATTTCTTAAGTATGAAGAGAAAGAC 1301

QY 301 ValValIleGluGluValGlnProTPAlaTyIleuSerCySArgIleHISValIleGlyTyThr 320
 Db 1302 GTTGTAGATGAAGAAACACATGAGGTATATCTAAATCGGCGCATATGATGATGAT 1361

QY 321 AsnTPArgIleAspSerAspThrGluAlaAsnGluValGlyCySProMetCySArgThr 340
 Db 1362 AACTGGAGGAACCAAGAGAAACGTGATGAGAAAGATCGTGAATGTCCTATGTGTATGCT 1421

QY 341 ValGlyProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAlaAlaGly 360
 DB 1422 GTTGGTCCCTATGCTCTCTGCTGGCTGATGGAAGCTGGATTATGTCAGCCGCCG 1481
 QY 361 ProProThHisAlaPheThrProCysGlyLysIleValCysSerGluLysSerAlaLysTyr 380
 DB 1482 CTTCCAAACCAACGCTTTAGCCGCTGGGATGTTGTTCCAGAAAGAACACTGCCCTAT 1541
 QY 381 TrpSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
 DB 1542 TGGTCCCAATCCCACTTCTCTATGTAACATCACTTTTCATGACGCTGCTCTTTGT 1601
 QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLeuLeuIlePheGlnGlyProIleAsp 420
 DB 1602 GCACATCAGTTGGCTGCTGTGACAAAGCCTACATCAGACTTATTTTCAAGGACCTCTAGAC 1661
 RESULT 15
 ADF83102
 ID ADF83102 standard; cDNA; 7136 BP.
 AC ADF83102;
 XX
 DT 26-FEB-2004 (first entry)
 DE Human Pellino homologue 1 gene, overexpressed in cancer.
 XX
 KM Human; Pellino homologue 1; Pellino1; cancer; cytosolic; vaccine;
 XX gene therapy; gene; oncogene; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 4039..5295
 FT /*tag= a
 XX
 PN MO2003100000-A2.
 PD 04-DEC-2003.
 XX
 PF 22-MAY-2003; 2003WO-US016049.
 XX
 PR 24-MAY-2002; 2002US-0382606P.
 PR 25-JUL-2002; 2002US-0398099P.
 XX
 PA (TULA-) TULARIK INC.
 XX
 PI Li J, Mu D, Yang J;
 XX
 DR WPI, 2004-035118/03.
 DR P-PSDB; ADF83103.
 DR GENBANK; NM_020651.
 XX
 PT diagnosing a cancer in a mammal comprises determining RecQL5, CTXL,
 PT USP13, MCL1, or Pellino 1 gene copy number in a biological sample from a
 PT region of the mammal that is suspected to be precancerous or cancerous.
 PS
 PS Claim 179; SEQ ID NO 9; 174pp; English.
 XX
 CC The present sequence is that of the gene encoding human Pellino homologue
 CC 1 (Pellino1), a protein associated with the kinase domain of activated
 CC Pelle. The invention is based on the finding of the overexpression of
 CC Pellino1 and other genes (RecQL5, CTXL, USP13 and MCL1) in certain
 CC cancers, including breast cancer, colon cancer, lung cancer and ovarian
 CC cancer, and the frequent amplification of these genes in cancer cells.
 CC The genes, and their expression products, can be used diagnostically or
 CC as targets for cancer therapy. They can also be used to identify and
 CC design compounds useful in the diagnosis, prevention and therapy of
 CC tumours and cancers, in vaccine development, and in methods for
 CC determining the efficacy of a treatment regime. A claimed method for
 CC inhibiting cancer or precancerous growth, especially in ovarian or lung
 CC tissue, uses an inhibitor that interacts with Pellino1 DNA or RNA. The
 CC inhibitor is a small interfering RNA (siRNA), microRNA (miRNA), an
 CC antisense RNA, and antisense DNA, a decoy molecule, a decoy DNA, a

CC ribozyme or small molecule.
 XX
 SQ Sequence 7136 BP; 1945 A; 1333 C; 1399 G; 2459 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 86-165 Length: 7136
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.48% Conservative: 38
 Best Local Similarity: 81.43% Mismatches: 38
 Query Match: 83.71% Indels: 2
 DB: 12 Gaps: 1
 US-10-041-030-4 (1-420) X ADF83102 (1-7136)
 QY 1 MetPheSerProGlyGlnGluGlnHisCysValaProAsnLysGluProValIleTyrGly 20
 DB 4039 ATGTTTCTCCGTGATCAAGAAATATAT-----CATCTAAAGCAACAGTAAATATGCT 4092
 QY 21 GluLeuValIleLeuGlyTyrTranGlyAlaLeuProAsnGlyAspArgGlyTranArgLys 40
 DB 4093 GAACCTATTGCTTATGGCTATATATGGCTCTCCAAATGGCGATAGAGAAAGAGAGAA 4152
 QY 41 SerArgPheAlaLeuTyrIleAspArgProLysAlaAsnGlyValIleAspSerThrValHis 60
 DB 4153 AGTAGGTTGCTTTGTTTAAAGACCTAAGCAATGGGTGAAGCCAGCATGTGCAT 4212
 QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysValGlyGlnHisSerIleSer 80
 DB 4213 ATTGCTTGACTCCTCAGCTGCAAGCAATAGCAACAAAGACAGCATATGATATCA 4272
 QY 81 TyrThrLeuSerArgAsnGlnThrValValIleGluTyrThrHisAspLysAspThrAsp 100
 DB 4273 TATACTTGTCTCGGGCCAGCATGTGGTGTGTAATATATCATGACAGACATGAT 4332
 QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValIleThrAspThrIle 120
 DB 4333 ATGTTTCAATTTGGCCGGCTGACTGAAGCCCATGATTTTGTATATCTACACGGTT 4392
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 4393 CCTGGAAAGTCAAAAGTATCTGATATACAGTACAGTCAAAAGCATATATCAAGATTGGCC 4452
 QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheHisAlaGlyPhe 160
 DB 4453 TGCAGATATCATATGAGAACGGAATCTCCCTTACAGCAGGATTTATCTGCAGGATTT 4512
 QY 161 AspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTyrLysAsnProAspGly 180
 DB 4513 GACTCATCAAAAACATCTTTCTTGGGAGAAAGCTGCCAAATGGAAAGACATCAGATGGA 4572
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 DB 4573 CAGATGATGGCTTGACCACTAATGATGTTCTTGTGATGATCATCCAGCAATGGGTTCA 4632
 QY 201 GluGlnSerGlnProGlyValIleTranGlyLysIleSerValCysGlyAspValIleThrLeu 220
 DB 4633 GAAGATCCCAAGCCTGGAATATGAGAAATATCGTGTGTGGAATATATTTATGACCTA 4692
 QY 221 ArgGluThrArgSerAlaGlnGlnIleArgGlyLysLeuValGluSerGluThrAsnValLeu 240
 DB 4693 CTTGAAACCAAGATCCGCTCAGCAGAGAGAAATGTTGAAATTCAAATCATCGTTA 4752
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrIleLeuLeuTrpArgThrAlaAspGly 260
 DB 4753 CAAGATGGCTCGTTATTAATCACTCTGTGTGCAACATTTGTATGGGCTCTCAGAAAGCC 4812
 QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnLysLeuAlaHis 280
 DB 4813 CTTTCCCAACATCTTCAACCTGAGCACTTTTAAAGCATTTTAAAGCAGAAATCAATGACGA 4872
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
 DB 4873 CAGCTCATGCTGCTGTAGGTTCAACATAGCATTTCTTAATGTAAGAGAGAAAGAC 4932

```
QY 301 ValValGluGluLysGlnProTrrpAlaTyrLeuSerCysGlyYHIsValHIsGlyTyrHis 320
DB 4933 GTTGTAGATGAAACAAACCAAGGATATCTAAACTGCGGCCAGTACATGACTATCAT 4992
QY 321 AenTrrpGlyHIsArgSerAspThrGluIaAenGluArgGluCysAspMetCysArgThr 340
DB 4993 AACTGGGGAACAAAGAAAGAGTATGGAAGAAAGATCGGAATGTCCTATGTGTAGGTCT 5052
QY 341 ValGlyProTyrValProLeuTrrpLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
DB 5053 GTTGGTCCCTATGTTCTCTGTGGCTTGATGTGAAGCTGGAATTTATGTGACGCGCGC 5112
QY 361 ProProThrHIsAlaPheThrProCysGlyYHIsValCysSerGluLysSerAlaLysTyr 380
DB 5113 CTTCCAAACCAATGCCGTTTAGCCCGTGTGGCAATGTGTTCAGAAAGCAAACTGCTAT 5172
QY 381 TrpSerGlnIleProLeuProHIsGlyThrHIsAlaPheHIsAlaAlaCysPropheCys 400
DB 5173 TGGTCCCAAGATCCCACTTCTCATGTACTCATACTTTTCATGCAGCCGTCCCTTTTGT 5232
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
DB 5233 GCACATCAAGTGGCTGTGTGAACAAGGCTACATCAGACTTATTTTCAAGACCTTAGAC 5292
```

Search completed: December 9, 2004, 13:47:55
Job time : 820 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2004, 15:53:42 ; Search time 845 Seconds
(without alignments)
2740.448 Million cell updates/sec

Title: US-10-041-030-4
Sequence: 1 NPSPOBHCAPNKEPVKYG.....ATQVGEONCITLIFQGPID 420

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool.p/US10041030/runat 09122004 095519 7785/app.query.fasta_1.583
-DB=Published Applications NA -QMT=fastcap -SUFFIX=rmp -MIMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -INIT=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcio -NOR=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USBR=US10041030 @CGN 1.1 912 @runat 09122004 095519 7785
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FOAPOP=6 -FOAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2290	100.0	1263	15 US-10-085-117-288	Sequence 288, App
2	2290	100.0	5579	14 US-10-197-666A-135	Sequence 135, App
3	2290	100.0	5597	15 US-10-085-117-287	Sequence 287, App
4	2290	100.0	5921	13 US-10-041-030-3	Sequence 3, App
5	2283	99.7	1263	9 US-09-843-905A-7	Sequence 7, App
6	2283	99.7	1263	9 US-10-317-250-7	Sequence 7, App
7	2283	99.7	1263	16 US-10-258-703-7	Sequence 7, App
8	2175.5	95.0	1260	9 US-09-843-905A-5	Sequence 5, App
9	2175.5	95.0	1260	15 US-10-317-250-5	Sequence 5, App
10	2175.5	95.0	1260	15 US-10-085-117-285	Sequence 285, App
11	2175.5	95.0	1260	16 US-10-258-703-5	Sequence 5, App
12	2111.5	92.2	1717	14 US-10-197-666A-133	Sequence 133, App
13	2111.5	92.2	1734	15 US-10-085-117-284	Sequence 284, App
14	1917	83.7	1257	9 US-09-843-905A-3	Sequence 3, App
15	1917	83.7	1257	15 US-10-317-250-3	Sequence 3, App
16	1917	83.7	1257	16 US-10-258-703-3	Sequence 3, App
17	1917	83.7	7136	13 US-10-041-030-1	Sequence 1, App
18	1917	83.7	7136	16 US-10-443-108-9	Sequence 1, App
19	1912	83.5	1257	15 US-09-843-905A-1	Sequence 1, App
20	1912	83.5	1257	15 US-10-317-250-1	Sequence 1, App
21	1912	83.5	1257	16 US-10-258-703-1	Sequence 1, App
22	1642	71.7	1338	9 US-09-843-905A-11	Sequence 11, App
23	1642	71.7	1338	15 US-10-317-250-11	Sequence 11, App
24	1642	71.7	1338	16 US-10-258-703-11	Sequence 11, App
25	1482.5	64.7	2508	15 US-10-104-047-1958	Sequence 1958, App
26	1064	46.5	202802	15 US-10-085-117-286	Sequence 286, App
27	1048	45.8	155579	15 US-10-085-117-283	Sequence 283, App
28	1022	44.6	541	15 US-10-029-386-22982	Sequence 22982, App
29	974	42.5	528	15 US-10-029-386-20504	Sequence 20504, App
30	808	35.3	24173	10 US-09-764-891-9294	Sequence 9294, App
31	808	35.3	24173	14 US-10-091-572-855	Sequence 855, App
32	775	33.8	465	9 US-09-867-550-1719	Sequence 1719, App
33	763	33.3	514	15 US-10-029-386-6790	Sequence 6790, App
34	763	33.3	514	15 US-10-029-386-9281	Sequence 9281, App
35	717	31.3	468	10 US-09-918-995-15544	Sequence 15544, App
36	678	29.6	377	9 US-09-867-550-1719	Sequence 1719, App
37	516	22.5	812	10 US-09-764-891-7046	Sequence 7046, App
38	516	22.5	813	10 US-09-764-891-7044	Sequence 7044, App
39	516	22.5	813	10 US-09-764-891-7047	Sequence 7047, App
40	222.5	9.7	196	10 US-09-764-891-9297	Sequence 9297, App
41	222.5	9.7	196	14 US-10-091-572-858	Sequence 858, App
42	171	7.5	5238	13 US-10-027-632-76172	Sequence 76172, App
43	171	7.5	5238	15 US-10-027-632-76172	Sequence 76172, App
44	170	7.4	1452	10 US-09-764-891-9295	Sequence 9295, App
45	170	7.4	1452	14 US-10-091-572-856	Sequence 856, App

ALIGNMENTS

RESULT 1
US-10-085-117-288
Sequence 288, Application US/10085117
Publication No. US2003023234A1
GENERAL INFORMATION:
APPLICANT: Englehard, David
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: 529,526/0121
CURRENT APPLICATION NUMBER: US/10/085,117
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: US 09/798,586
NUMBER OF SEQ IDS: 361
SOFTWARE: Padded for windows Version 4.0
SEQ ID NO 288
LENGTH: 1263
TYPE: DNA
ORGANISM: Homo sapiens
US-10-085-117-288

Alignment Scores:

Pred. No.: 2,556-259 Length: 1263
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 15 Gaps: 0

US-10-041-030-4 (1-420) x US-10-085-117-288 (1-1263)

```

QY      1 MetPheSerProGlyGlnGlnGlnUhiGcyValAProAenlyGluProVallyGly 20
DB      1 AAGTTTCCCTGGGCGAGAGAACACTGGCCCCCAATAAGAGCAGTGAATAACGGG 60
QY      21 GluLeuValValLeuGlyTyrAenGlyAlaLeuProAenGlyAAspArgGlyArgGly 40
DB      61 GAGCGTGTGTCTGGGTACATAGTGTCTTACCCCAATGAGATAGAGAGAGAGAA 120
QY      41 SerArgPheAlaLeuTyrLySarArgProLySaLaAenGlyValLySProSerThrValHis 60
DB      121 AGTAGATTGGCTCTACAAAGCGGCCCAAGGCAATGGTCAAAACCCAGCACGTCAT 180
QY      61 ValLieserThrProGlnAlaSerLySaLaLieserCySlyGlyGlnHisSerLieser 80
DB      181 GTGAATATCACGCCCCAGGCATCCAAAGCTATCAGCTGCAAGAGTCAACACAGTATATCC 240
QY      81 TyrThrLieserArgAsnGlnThrValValGlnTyrThrHisAspLySaSphThrAp 100
DB      241 TACACTTTGTCAAGGAATACAGCTGTGTGGAGTACACATGATTAAGATACGGAT 300
QY      101 MetPheGlnValGlyArgSerThrGlySerProLiesPheValValThrAspThrIle 120
DB      301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCATACGACTTGCTGTGTCAACAGACGANT 360
QY      121 SerGlySerGlnAsnThrAepGlnAaglnLietrGlnSerThrIleSerArgPheAla 140
DB      361 TCTGGCAGCCAGAACACGAGCAGAACCCAGATCAACAGACACCATATCCAGTTCGCC 420
QY      141 CysArgGlyLeValCysAspArgAenGluProTyrThrAlaArgIlePheAlaGlyPhe 160
DB      421 TGCAGATCTGTGGCAGAGAAATGAACCTTACACAGACGGAATTTCCGCCCGGATTT 480
QY      161 AppSerSerLySaSnLlePheLeuGlyGlyLySaLaAlySlyTrpLySaSnProAspGly 180
DB      481 GACTCTTCCAAAACATATTTCTTGAGAAAAGCAGCAAAAGTGAAAAACCCGACGCG 540
QY      181 HisMetAepGlyLeuThrThrAenGlyValLeuValMetHisProArgGlyGlyPheThr 200
DB      541 CACATGGATGGGCTCACTAATAAGCGCTCTGTGTGATGATCCACAGAGGGGCTTTCACC 600
QY      201 GluGlySerGlnProGlyValTyrPArgGlnLieserValCysGlyAAspValTyrThrLeu 220
DB      601 GAGGATGCCAGCCCGGGGTCTGGGCGCAGATCTCTGTGTGAGATGTGTACACCTTG 660
QY      221 ArgGlnThrArgSerAlaGlnGlnArgGlyLyLyLeuValGlnSerGlnThrAsnValLeu 240
DB      661 CGAAGAAACAGGATCGGCGCCAGCAAGAGAAAGCTGTGTGAAGAGAGCAACGTCCTG 720
QY      241 GlnAepGlySerLeuLleAepLeuCySgLyAlaThrLeuLeuTyrPArgThrIleAepGly 260
DB      721 CAGGACGGCTCCCTCATTTGACTGTGTGGGCGCCTCTCTCGAAGAACACAGATGGG 780
QY      261 LeuPheHisThrProThrGlnLyHisIleGlnAlaLeuArgGlnGlnLleAsnAlaAla 280
DB      781 CTTTTCATCTCTCAACTCAGAACGACATAGAACCCCTCCGCAAGAGATTAAAGCGGCC 840
QY      281 ArgProGlnCySProValGlyLeuAsnThrLeuAlaPheProSerLleAsnArgLySgln 300
DB      841 CGGCTCAAGTGTCTGTGGGCTCAACACCTGGCTTCCAGCATCAACAGAAAGAG 900
QY      301 ValValGlnGlnLyGlnProTyrPalAlyTyrLeuSerCySgLyHisValHisGlyTyrHis 320
DB      901 GTGTGTGAGAGAACAGCCCTGTGGCATATATTCAGTTGTGGCCAGGTGACGGGTACAC 960
  
```

```

QY      321 AenTPGlyHisArgSerAspThrGlnAlaAenGluAArgGlnCySProMetCysArgThr 340
DB      961 AACTGGGCGCATCGAGTGAACCGAGGCGCAACAGAGAGAGTGTCCATGTGCGAGACT 1020
QY      341 ValGlyProTyrValProLeuTyrLeuGlyCySglnAlaGlyPheTyrValAspAlaGly 360
DB      1021 GTGGGCGCCCTATATGTCTCTGTGGCTGTGTGAGCAGATTTTATGTAGACGACAGA 1080
QY      361 ProProThrHisAlaPheThrProCySgLyHisValCysSerGlnLySserAlaSerTyr 380
DB      1081 CGGCAACTCATGTCTTACCTCCCTGAGACCGTGTCTGGAGAACTCTGCAAAATAC 1140
QY      381 TrpSerGlnLleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCySProPheCyS 400
DB      1141 TGTCTCAGATCCGCTTCTCATGGAATCATGCAATTTCAAGCTTGGCTTCTGTCTGT 1200
QY      401 AlaThrGlnLeuValGlyGlnGlnAenCySlyLeuLeuLlePheGlnGlyProLiesAp 420
DB      1201 GCTACACAGCTGTGTGGGAGCAAACTGCATCAAAATTAATTTCCAAAGTCCAATTGAC 1260

RESULT 2
US-10-197-666A-135
; Sequence 135, Application US/10197666A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAHI KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: Etk1 phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197, 666A
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 135
; LENGTH: 5579
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (177)..(1436)
US-10-197-666A-135

Alignment Scores:
Pred. No.: 2,176-258 Length: 5579
Score: 2290.00 Matches: 420
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-041-030-4 (1-420) x US-10-197-666A-135 (1-5579)
QY      1 MetPheSerProGlyGlnGlnGlnUhiGcyValAProAenlyGluProVallyGly 20
DB      177 AAGTTTCCCTGGGCGAGAGAAACACTGGCCCCCAATAAGAGCAGTGAATAACGGG 236
QY      21 GluLeuValValLeuGlyTyrAenGlyAlaLeuProAenGlyAAspArgGlyArgGly 40
DB      237 GAGCGTGTGTCTGGGTACATAGTGTCTTACCCCAATGAGATAGAGAGAGAGAA 296
QY      41 SerArgPheAlaLeuTyrLySarArgProLySaLaAenGlyValLySProSerThrValHis 60
  
```

Db 297 AGTAGATTGGCTCTTCAAGCGGCCGCAAGCAAAAGTGTCTCAAGCCAGCAGCTTCAT 356
 Qy 61 ValIleSerThrProGlnAlaSerIleValIleSerCysValyGlyGlnHisSerIleSer 80
 Db 357 GTGATATCCAGCGCCAGGCGATCCAAAGGCTATCAGCTGCAGAAAGTCAACAGTATATCC 416
 Qy 81 TyrThrLeuSerArgSerGlnThrValValValGluTyrThrHisAspIleAspThrAsp 100
 Db 417 TACACTTGTCAAGAAATCAGATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 476
 Qy 101 MetPheGlnValAlaIleValSerThrGlnSerProIleAspPheValValThrAspThrIle 120
 Db 477 ATGTTTCAGGTGGGCGAGATCAACAGAAAGCCATGACTCTGTTGTCACAGACAGATT 536
 Qy 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 537 TCTGGCAGCAGAAACAGGACGAAAGCCAGATCACAGAGCAGCATATCCAGTTCCGCC 596
 Qy 141 CysArgIleValCysAspArgSerGlnProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
 Db 597 TCGAGATGTGTGGCAGAGATGAACCTTACACAGCAGCGATATTCGCCCGCGGATTT 656
 Qy 161 AspSerSerIleAsnIlePheLeuGlyGluValAlaAlaValThrPheAspProAspGly 180
 Db 657 GACTCTTCCAAAACATATTTCTTGGAGAAAGGCGCAAGTGGAAAGCCCGACGGC 716
 Qy 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisAspArgGlyGlyPheThr 200
 Db 717 CACATGGATGGGCTCACTAATAGCGCTGCTGGTATGATCATCAGAGGGGCTTCAC 776
 Qy 201 GlnGluSerGlnProGlyValIleTpaArgGluIleSerValCysGlyAspValIleThrLeu 220
 Db 777 GAGAGATCCAGCGCGGGCTGGCGCGAGATCTGTCTGTGGAAATGTGACACCTTG 836
 Qy 221 ArgGluThrArgSerAlaGlnIleValGlyValLeuValGluSerGlnThrAsnValIle 240
 Db 837 CAGAAACAGAGTCCGCGCCAGCAGCAAGAAAGCTGTGGAAAGTGAAGCAACGTCCTG 896
 Qy 241 GlnAspGlySerLeuIleAspLeuValValAlaThrLeuLeuTpaArgThrAlaAspGly 260
 Db 897 CAGAGCGGCTCCCTCATTCAGCTGTGGGGCCACTCTCTCGAGAAACAGCAGATGG 956
 Qy 261 LeuPheHisThrProThrGlnIleValIleGluAlaLeuArgGlnIleValIleAla 280
 Db 957 CTTTTTCATATCTCAACTCAGAAAGCAGATGAAGCCCTCGCGCAGAGATTAACGCCGCC 1016
 Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgIleVal 300
 Db 1017 CGGCTCAATGTCCTGTGGGGCTCAACACCTGGCTTCCCGCATCAACAGAAAGAG 1076
 Qy 301 ValIleGluGluValGlnProTpaAlaTyrLeuSerCysGlyValIleValHisGlyTyrHis 320
 Db 1077 GTGGTGAAGAGAAAGACCGCTGGCATATCTCAGTTGTGGCCAGCGCAGCGGTACAC 1136
 Qy 321 AsnTyrGlyHisArgSerAspThrGlnAlaAsnGlnArgGluCysProMetCysAspGln 340
 Db 1137 AACTGGGGCATTCGAGTACAGAGAGCCCAAGAGAGAGAGTCCCATGTGCAGAGAT 1196
 Qy 341 ValIleProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAlaAspIle 360
 Db 1197 GTGGGCGCTTATGTGCTCTCTGTGGCTGTGAGGAGAGATTTATGTAGCGAGAG 1256
 Qy 361 ProProThrHisAlaPheThrProCysGlyValIleValCysSerGluIleSerAlaValTyr 380
 Db 1257 CGGCAACTCATGCTTCACTCCCTGTGACAGTGTGCTCGGAGAAAGTCTGCAAAATAC 1316
 Qy 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
 Db 1317 TGGTTCAGATCCCGTGTGCTCAAGAACTCAAGCATTCACGCTCTTGCCTTCTCT 1376
 Qy 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleValLeuIlePheGlnGlyProIleAsp 420

Db 1377 GCTACACAGCTGTGGGAGCAAGAAAGTGCATCAATTAATTTTCCAGGTCCAAATGAC 1436
 RESULT 3
 US-10-085-117-287
 ; Sequence 287, Application US/10085117
 ; Publication No. US2003023234A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Moritz, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
 ; FILE REFERENCE: 529452000121
 ; CURRENT APPLICATION NUMBER: US/10/085,117
 ; PRIOR FILING DATE: 2002-02-27
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; NUMBER OF SEQ ID NOS: 361
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 287
 ; LENGTH: 5597
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-085-117-287
 Alignment Scores:
 Pred. No.: 2,18e-258 Length: 5597
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 15 Gaps: 0
 US-10-041-030-4 (1-420) x US-10-085-117-287 (1-5597)
 Qy 1 MetPheSerProGlnGluGlnIleValCysValProAsnIleGluProValIleTyrGly 20
 Db 177 ATGTTTCCCGCGCGCAGAGAAACCTGCGCCCAATTAAGAGCAGTGAATACGGG 236
 Qy 21 GluLeuValIleGluIleTyrThrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGly 40
 Db 237 GAGCTGGTGGTCTCGGTACATAGTGTCTTACCCAAAGAGATAGAGAGAGAGAA 296
 Qy 41 SerArgPheAlaLeuTyrTyrAspProIleValAlaAsnGlyValIleProSerThrValHis 60
 Db 297 AGTAGATTGGCTCTTCAAGCGGCCGCAAGCAAAAGTGTCTCAAGCCAGCAGCTTCAT 356
 Qy 61 ValIleSerThrProGlnAlaSerIleValIleSerCysValyGlyGlnHisSerIleSer 80
 Db 357 GTGATATCCAGCGCCAGGCGATCCAAAGGCTATCAGCTGCAGAAAGTCAACAGTATATCC 416
 Qy 81 TyrThrLeuSerArgSerGlnThrValValValGluTyrThrHisAspIleAspThrAsp 100
 Db 417 TACACTTGTCAAGAAATCAGATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 476
 Qy 101 MetPheGlnValAlaIleValSerThrGlnSerProIleAspPheValValThrAspThrIle 120
 Db 477 ATGTTTCAGGTGGGCGAGATCAACAGAAAGCCATGACTCTGTTGTCACAGACAGATT 536
 Qy 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 537 TCTGGCAGCAGAAACAGGACGAAAGCCAGATCACAGAGCAGCATATCCAGTTCCGCC 596
 Qy 141 CysArgIleValCysAspArgSerGlnProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
 Db 597 TCGAGATGTGTGGCAGAGATGAACCTTACACAGCAGCGATATTCGCCCGCGGATTT 656
 Qy 161 AspSerSerIleAsnIlePheLeuGlyGluValAlaAlaValThrPheAspProAspGly 180
 Db 657 GACTCTTCCAAAACATATTTCTTGGAGAAAGGCGCAAGTGGAAAGCCCGACGGC 716
 Qy 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisAspArgGlyGlyPheThr 200
 Db 717 CACATGGATGGGCTCACTAATAGCGCTGCTGGTATGATCATCAGAGGGGCTTCAC 776

```

QY 201 GIUGLUSERG1PROG1YVALTPARGLN1LESERVALCYSG1YASPVATYTRThleu 220
DB 777 GAGGAGTCCCAAGCCCGGGGTCTGGCCGAGATCTCTGTCTGGAGATGTGACCTTG 836
QY 221 ARGGLUTHARGSER1AG1NG1NARG1LYLVLEUVALGIUSERG1UTHRANVALLEU 240
DB 837 CGAGAAACGAGGTCCGCCAGCAACGAGGAAAGCTGTGGAAAGTGAAGCAACGCTCG 896
QY 241 G1NAPRG1YSER1LEU1LEAPLEUCYSG1YALATHLEULEUTPRARGH1LAASPG1Y 260
DB 897 CAGGAGGCTCCCTCACTTACCTGTGTGGGCGCACTCTCTCTGAGAAACAGCAATGGG 956
QY 261 LEUPHEH1ETHPROTHRG1NLYNH11EG1UALALEUARGLNGL1LEASNA1A1A 280
DB 957 CTTTTTCATCTCACTCAAGTCAAGAAAGCATAGAAAGCCCTCCGCGAGAGATTAAACGCCG 1016
QY 281 ARGPROG1NCYARPROVALG1YLEUANTH1LEUALAPHROSER1LEASNARG1YSG1U 300
DB 1017 CGGCGCTCAGTGTCTGTGGGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 1076
QY 301 VAL1VALGIUGLULYRG1NPROTPRALATYR1LEUSERCYSG1YH1EVALH1SG1YTRH1S 320
DB 1077 GTGGTGAGAGAAAGAGAGCCCTGGGCAATCTCAGTTGTGGCCACGTGCAAGGTAACAC 1136
QY 321 AANTPRG1YH1ARGSER1APTHRG1UALAANG1UARG1UCYARPROK1CYASARG1THR 340
DB 1137 AACTGGGGGCATCGAGATGACACGAGGCGCAACGAGAGAGTCTCCCATGTGCGAGACT 1196
QY 341 VAL1GYPROTYRVALPROLEUTPR1LEUG1YCYSG1UAL1G1YPHETRYVAL1APALAG1Y 360
DB 1197 GTGGGGCCCTATGTGCTCTCTGTGGCTGTGGCGAGAGATTTATGTACACGACAGA 1256
QY 361 PROPTOTH1H1ALPHETHRG1PROCYSG1YH1SVALCYASERGLN1YASER1ALVSTYR 380
DB 1257 CCGCAACTCAAGCTTCTCACTCTGTGGAGACAGTGTGCTCGAGAGAGTCTGCAAAATAC 1316
QY 381 TRPSERGLN1LEPROLEUPROH1SG1YTRH1SALAPHH1E1A1A1A1CYSPROPHCY 400
DB 1317 TGGTTCAGATCCCGTGTGCTCATGSAATCATGATTCATGAGTGTGCTTGTCTGTCTGT 1376
QY 401 ALATHRG1NLEUVALG1YGLUG1NANCYV11ELYVLEU1LEPHG1NGLYPRO1LEAP 420
DB 1377 GCTACACAGCTGTGGGGAGAGAACTGATCAATTAATTTTCCAAAGTCAATTGAC 1436

RESULT 4
US-10-041-030-3
; Sequence 3, Application US/10041030
; Publication No. US20020150934A1
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Mu, David
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
; FILE OF INVENTION: Pellino Polypeptides and Polynucleotides
; FILE REFERENCE: 018781-006810US
; CURRENT APPLICATION NUMBER: US/10/041,030
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/259,502
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5921
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(1763)
; OTHER INFORMATION: human pellino 2
US-10-041-030-3

```

Alignment Scores:

Pred. No.:	2,376-258	Length:	5921
Score:	2290.00	Matches:	420
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-10-041-030-4 (1-420) x US-10-041-030-3 (1-5921)

```

QY 1 MetPheSerProG1YGLNGLN1H1SCYSA1APROANLYRG1UPROVAL1YSTRG1Y 20
DB 501 ATGTTTCCCTCCGCGCAAGAGAAACATGCGCCCAATAAGAGCCAGTGAATAACCGG 560
QY 21 GIU1EUV1VAL1LEUG1YTRARG1YVAL1LEUPROANG1YASPVATYTRARG1Y 40
DB 561 GAGCTGGTGCTCGGTACATGCTCTTACCAATGAGATAGAGACGAGGAGAA 620
QY 41 SERARGPHEALALEUTRYLYEARPROLYSAL1AANG1YVAL1LYSPROSERTHRYALH1S 60
DB 621 AGTAGATTTCCTCTTACAGCGGCCCAAGCAATGCTCAACCAAGACCGTCCAT 680
QY 61 VAL1LESERTHRG1NAL1SER1YAL1LESERCYSLYRG1YGLNH1SER1LESER 80
DB 681 GGTATATCCAGCGCCCAAGGATCCAAAGGCTATCAGCTCAAGGTCAACAGATATATCC 740
QY 81 TYRTH1LEUSERARG1ANG1NTHRYVAL1VAL1GIUYTRTH1ASPVLYASPVTHRAP 100
DB 741 TACACTTGTCAAGAAATCAGCTGTGGTGAGTACACATGATTAAGATTCGAGT 800
QY 101 METPHEG1NVALG1YARGSERTHRG1USERPRO1LEAPHEVAL1YTHRAPTH1LE 120
DB 801 ATGTTTCAAGTGGGAGATCAACAGAAAGCCCTATGACCTTGTTGACAGACAGATT 860
QY 121 SERG1YSERGLNANTH1RASPGLUAL1AG1N1LETHRG1N1SERTH1LESERARGPHEAL 140
DB 861 TCTGGCACACGAGAACCGAGAGCCAGATGACACAGACACCATATCCAGGTTCCGC 920
QY 141 CYSAAG11EVALCYASPARARG1UPROTYRTH1AARG11EPHEAL1A1AG1YPHE 160
DB 921 TGCAGGATCGTGTGGAGAGAGATGACCTTACACAGACGATATTCGCGCCGCAATT 980
QY 981 AEPSE1ERYASN11LEPHELEUG1YGL1YVAL1A1Y1STR1LYASNP1ROASPG1Y 180
DB 1041 CACATGAGATGGGCTCACTTAATGCGTCTGTGGATGACATCCAGAGGGGGCTTTCAC 1100
QY 201 GIUGLUSERG1NPROG1YVALTPARGLN1LESERVALCYSG1YASPVATYTRThleu 220
DB 1101 GAGAGTCCCAAGCCCGGGGTCTGGCGAGATCTGTCTGTGGAGATGTGACCTTG 1160
QY 221 ARGGLUTHARGSER1AG1NG1NARG1LYLVLEUVALGIUSERG1UTHRANVALLEU 240
DB 1161 CGAGAAACGAGGTCCGCCAGCAACGAGGAAAGCTGTGGAAAGTGAAGACCAACGCTCG 1220
QY 241 G1NAPRG1YSER1LEU1LEAPLEUCYSG1YALATHLEULEUTPRARGH1LAASPG1Y 260
DB 1221 CAGGAGGCTCCCTCATATGAGCTGTGTGGGCGCACTCTCTCTGGAACACGCAATGG 1280
QY 261 LEUPHEH1ETHPROTHRG1NLYNH11EG1UALALEUARGLNGL1LEASNA1A1A 280
DB 1281 CTTTTTCATCTCACTCAAGTCAAGAAAGCATAGAAAGCCCTCCGCGAGAGATTAAACGCCG 1340
QY 281 ARGPROG1NCYARPROVALG1YLEUANTH1LEUALAPHROSER1LEASNARG1YSG1U 300
DB 1341 CGGCTCAGTGTCTGTGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 1400
QY 301 VAL1VALGIUGLULYRG1NPROTPRALATYR1LEUSERCYSG1YH1EVALH1SG1YTRH1S 320
DB 1401 GTGGTGAGAGAAAGAGAGCCCTGGGCAATCTCAGTTGTGGCCAGTGCACGCGGTACAC 1460

```



```

QY 321 AERTTGLYHIAERSEARPTHTGUAIAAENGUAIGUCYSPROMETCYBATGTHR 340
DB 1461 AACTGGGGCCATCGGATGACACGAGGCAAGAGGAGTGTCCATGTGCAGAGACT 1520
QY 341 VALGLPRTYRVALPRTLEUTRPLEUGLYCYSGUAAAGLYPHETYRVALAEPALAGLY 360
DB 1521 GTGGGCCCCCTATGTGCTCTCTGCTTGGCTTGAGGACAGATTATATGACGAGGA 1580
QY 361 PROPTHTHIAEAEPHEPTHTPCYSGLYHIAVALCYSESGILUYSESERIALUYSTYR 380
DB 1581 CGGCCAATCATGCTTCACTCCCTGTGACACGTGTGCTCGGAAAGTCTGCAAAATAC 1640
QY 381 TRPSERGINLEPRTLEUPROHISGLYTHRHIAEAEPHEHIAAALACYPRTOPHECYG 400
DB 1641 TGGTTCAGATCCCGTGTGCTCATGGAACATGACATTCACCTGCTGCTTCTGT 1700
QY 401 ALATHGLINLEVALGLYGLUGLINAENCYVILAEVLEUILEPHEGINGLYPROILEAP 420
DB 1701 GCTACACAGCTGTGGGAGCAAACTGCATCAAAATTAATTTCCAAAGGTCCAAATTGAC 1760

RESULT 5
US-09-843-905A-7
; Sequence 7, Application US/09843905A
; Patent No. US20020168683A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-843-905A-7

Alignment Scores:
Pred. No.: 1,7e+258 Length: 1263
Score: 2283.00 Matches: 419
Percent Similarity: 99.764 Conservative: 0
Best Local Similarity: 99.764 Mismatches: 1
Query Match: 99.694 Indels: 0
Gaps: 0
DB: 9

US-10-041-030-4 (1-420) x US-09-843-905A-7 (1-1263)
QY 1 MetPhSerProGlynglugluhiahCyAlaProAmlysluProVallybTYRGLY 20
DB 1 AAGTTTCCCTCGCCAGAGAACTGCGCCCAATTAAGAGCAAGTGAATACCGG 60
QY 21 GLULEVALVALLEUGLYTYRAENGILYALAEUPROAENGILYAEPTRGILYARGAGLY 40
DB 61 GAGCTGTGTGCTCGGATCAATGTGTCTTACCAATGCAATGAGAGACGAGAGAAA 120
QY 41 SEARPHALAEUPTLYSARGPROLYSALAEAGLYVALIYSPROSEPTHTYRVALHIS 60
DB 121 AGTAGATTGCTCTTCAAGCGGCCCAAGCAAAATGTGTCAAAACCCAGCACCGTCCAT 180
QY 61 VALIISERTHTPROGILNALSERTYSAIILESECYSELYSGILYGLINHISERTIISER 80
DB 181 GTGATATTCACGCGCCACGAGATCCAAAGGCTATCACTGCAAAAGTCAACACATATATCC 240
QY 81 TYRTHLEUSERARGANGINHRVALVALVALGLYTYRTHHIAEPLYSAEPTHTAEP 100
DB 241 TACACTTTGCAAGGAATCAAGACTGTGTGTGAGATACACACATATTAAGATACGAGAT 300
QY 101 MEPTHEGLINLEVALYRSEPTHTRGUSERPROILEAPHEVALVALTHRAEPTHTRIE 120

```

```

DB 301 AATGTTCAAGTGGGAGATGATCAAGAAAGCCCTATGATGATTTGTTTCACAGACAGATT 360
QY 121 SERGLYSESGILNASTHTASPTGLUAGINILERTHINSETHRIIESERARGPHEAL 140
DB 361 TGTGGCAGCAAAACACGAGACGAGCCAGATCAACAGACCAATATCCAGGTTCCGC 420
QY 141 CYEARGILEVALCYEAPARGANGILUPRTYRTHRALAAGTILEPHEALAAAGLYPHE 160
DB 421 TCGAGGATGTGTGCGACGAGATGAACCTTACACAGCACGAGATATTCGCCCGGATTT 480
QY 161 APPSERSELYSAANILEPHELEUGLYULYSAIAALAEVLEPTRLYSANPRTAEPGLY 180
DB 481 GACTCTTCAAAACATATTTCTTGAGTAAGGACGCAAAAGTGGAAAAACCCGACGGC 540
QY 181 HISMEASPGLYLEUTHRTHRASNGILYALAEVALMECHIEPROARGLYGLYPHETHR 200
DB 541 CACATGATGGGCTCATCTAATGCGCGTGGTGAATGATCATCCACGAGGGGCTTCAAC 600
QY 201 GLUGLUSERGINPROGLYVALITPAARGIULIISERVALCYSGIYASPVALITYRTHLEU 220
DB 601 GAGAGTCCACAGCCCGGGGTCTGGCGGAGATCTGTCTGTGAGATGTATACACCTTG 660
QY 221 ARGGLUTHTARISERXALAGINGINARGGLYLYSELEVALGLUSESGIUTHRASNVALL 240
DB 661 CAGAAACCAAGTCCGACCAAGCAAGGAAAGCTGTGAAAGTGAACCAACGTTCTTG 720
QY 241 GINAEPLYSERLEULEAPLEUCYSGIYALATHLEULEUTRPAARGTHRAAEPGLY 260
DB 721 CAGGACGGCTCCCTATTATGACCTGTGTGGGCACTCTCTCGAGGAACAGACAGATGG 780
QY 261 LEUPHEHISTHTPROTHRGILNLSHISILIEGLUALAEUARGINGILULEANALAA 280
DB 781 CTTTTTCATATCTCAACTCAGAAACACATAGAAAGCCCTCCGACAGAGATTAAACCGCC 840
QY 281 ARSPROGILNCPROVALGLYLEUENHTRLAEUALAPHEPROSEPTHTAEANRGLYSGIU 300
DB 841 CGGCTCATGTCTCTGTGGGAGCTCAACACCTGCTTCCAGCATCAACAGAAAGAG 960
QY 301 VALVALGLUGLULYSEGINPTROTPTALATYRLEUSERCYSGIYHIAVALHISGLYTHR 320
DB 901 GTGGTGAAGAGAAAGACGCTGGGCAATATCTCAGTGTGTGGCAAGTGCACGGGTACAC 960
QY 321 AERTTGLYHIAERSEARPTHTGUAIAAENGUAIGUCYSPROMETCYBATGTHR 340
DB 961 AACTGGGGCCATCGGATGACACGAGGCAAGAGGAGTGTCCATGTGCAGAGACT 1020
QY 341 VALGLPRTYRVALPRTLEUTRPLEUGLYCYSGUAAAGLYPHETYRVALAEPALAGLY 360
DB 1021 GTGGGCCCTATGTGCTCTCTGCTTGGCTGTGAGGACGAAATTTATATGACGAGGA 1080
QY 361 PROPTHTHIAEAEPHEPTHTPCYSGIYHIAVALCYSESGILUYSESERIALUYSTYR 380
DB 1081 CGGCCAATCATGCTTCACTCCCTGTGACACGTGTGCTCGGAAAGTCTGCAAAATAC 1140
QY 381 TRPSERGINLEPRTLEUPROHISGLYTHRHIAEAEPHEHIAAALACYPRTOPHECYG 400
DB 1141 TGGTTCAGATCCCGTGTGCTCATGGAACATGACATTCACGCTGCTGCTTCTGT 1200
QY 401 ALATHGLINLEVALGLYGLUGLINAENCYVILAEVLEUILEPHEGINGLYPROILEAP 420
DB 1201 GCTACACAGCTGTGGGAGCAAACTGCATCAAAATTAATTTCCAAAGGTCCAAATTGAC 1260

RESULT 6
US-10-317-250-7
; Sequence 7, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; APPLICANT: Li, Xiaoxia
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-B

```

```
/ CURRENT APPLICATION NUMBER: US/10/317,250
/ CURRENT FILING DATE: 2002-12-11
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 7
/ LENGTH: 1263
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-317-250-7

Alignment Scores:
Pred. No.: 1,7e-258 Length: 1263
Score: 2283.00 Matches: 419
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 99.69% Indels: 0
DB: 15 Gaps: 0

US-10-041-030-4 (1-420) x US-10-317-250-7 (1-1263)

QY 1 MetPheSerProGluGlnGluGlnHisCyAlaProAlaSerGluProValIleTyrGly 20
DB 1 ATGTTTCCCTGGCCAGAGAACTGCGCCCAATTAAGAGCCAGTGAATACGGG 60
QY 21 GluLeuValIleuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgGly 40
DB 61 GAGCTGTGTGCTCGGAGTCAATGCTGTTTACCAATGAGATGAGAGACGAGAGAA 120
QY 41 SerArgPheAlaLeuTyrIleAspArgProValIleAsnGlyValIlePheSerThrValHis 60
DB 121 AGTAGATTGCTCTTACAAAGCGCCCAAGGCAATGGTGTCMAACCCAGCACCTGCAT 180
QY 61 ValIleSerThrProGlnAlaSerIleValIleSerCysValGlyGlnHisSerIleSer 80
DB 181 GTGATATCCAGCCCGCCAGGATCCAGAGCTTATCAGTGAAGGTCAACACATATATCC 240
QY 81 TyrThrLeuSerArgAsnGlnThrValIleValIleGluTyrThrHisAspIleAspThrAsp 100
DB 241 TACACTTTGTCAAGAAATCAGACTGTGTGTGTGAGATCAACATGATTAAGATACGAT 300
QY 101 MetPheGlnValIleValArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120
DB 301 ATGTTTCAAGTGGCGGATCAACAGAAACCCATTCGACTGTGTGTCAAGACACGATT 360
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 361 TCTGCGACGCAAAACCGAGCAAGCCCGAGATCAACAAGACCAATATCCAGTTCCGC 420
QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleArgIlePhe 160
DB 421 TCGAGGATCTGTGTGCGAGATGAACTTACACAGCAGGATATTGCGCCCGGATTT 480
QY 161 AspSerSerIleAsnIlePheLeuGlyGlyValAlaIleTyrIleAsnProAspGly 180
DB 481 GACTCTTCCAAAACATATTTCTTGAAGTAAAGCGCAAGGTGAAAAACCCCGACGCG 540
QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleValMetHisProArgGlyGlyPheThr 200
DB 541 CACATGAGTGGGCTCATCTAATGCGCTCTGTGTATGATCAATCCAGAGGGGCTTCACC 600
QY 201 GluGluSerGlnProGlyValIleTyrArgGlnIleSerValCysGlyAspValTyrThrLeu 220
DB 601 GAGGAGTCCCAAGCCCGGGGTCTGGCGGAGATCTGTGTGTGAGAGTGTATACCTTG 660
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyValIleValGluSerGluThrAsnValLeu 240
DB 661 CAGAGAAACAGAGTCCGCGCCAGCAACAGAGAAAGCTGTGAAATGAGACCAACGCTCG 720
QY 241 GlnAspGlySerIleuIleAspLeuCysGlyAlaThrLeuLeuTyrArgThrIleAspGly 260
DB 721 CAGGAGGCTCTCTATTAAGTCTGTGTGTGGGCACTCTCTCTGAGAAACAGACAGTGGG 780
QY 261 LeuPheHisThrProThrGlnIleHisIleGluAlaLeuArgGlnIleAsnAlaIle 280
```

```
|||||
DB 781 CTTTTCATATCTCCAACTCAGAAAGCATTAAGACCTCCGCGAGAGATTAAACCCGCC 840
QY 281 ArgProGlnCysProValIleuAsnThrLeuAlaPheProSerIleAsnArgIleGlu 300
DB 841 CGGCTCAGTGTCTGTGTGGGCTCAACACCTGGCGCTTCCCAAGATCAACAGAGAAAG 900
QY 301 ValIleGluGlyValProTyrPalaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
DB 901 GTGTGAGAGAAACAGCCCTGGGCATATCTCAATTGTGGCCACGTGACGGATCCAC 960
QY 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
DB 961 AACTGGGGCCATCGAGTGAACACGAGGCCAACAGAGAGAGTGTCCCATGTGCAGACT 1020
QY 341 ValGlyProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
DB 1021 GTGGGCCCCCTAATGTCCTCTGTGGCTGTGTGAGGAGATTTATGTAGACCGAGGA 1080
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluIleSerAlaIleTyr 380
DB 1081 CGGCCAACTCATGCTTTTCACTCTCTGTGACACGCTGTCTCGAGAAAGTCTGCATAATAC 1140
QY 381 TTPSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
DB 1141 TGGTCTCAGATCCGTTGCTTCATGAACTCATGCACTTACGCGTGTGCTTCTGT 1200
QY 401 AlaThrGlnLeuValIleGluGlnAsnGlyIleIleValIlePheGlnGlyProIleAsp 420
DB 1201 GCTACACAGCTGTGTGGGAGCAACATGCATCAATTAATTTCCAAAGTTCCAATTGAC 1260

RESULT 7
US-10-258-703-7
/ Sequence 7, Application US/10258703
/ Publication No. US20040034199A1
/ GENERAL INFORMATION:
/ APPLICANT: Bird, Timothy A.
/ APPLICANT: Cosman, David J.
/ TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
/ FILE REFERENCE: 2990-US
/ CURRENT APPLICATION NUMBER: US/10/258,703
/ CURRENT FILING DATE: 2002-10-24
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 7
/ LENGTH: 1263
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-258-703-7

Alignment Scores:
Pred. No.: 1,7e-258 Length: 1263
Score: 2283.00 Matches: 419
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 99.69% Indels: 0
DB: 16 Gaps: 0

US-10-041-030-4 (1-420) x US-10-258-703-7 (1-1263)

QY 1 MetPheSerProGluGlnGluGlnHisCyAlaProAlaSerGluProValIleTyrGly 20
DB 1 ATGTTTCCCTGGCCAGAGAACTGCGCCCAATTAAGAGCCAGTGAATACGGG 60
QY 21 GluLeuValIleuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgGly 40
DB 61 GAGCTGTGTGCTCGGAGTCAATGCTGTTTACCAATGAGATGAGAGACGAGAGAA 120
QY 41 SerArgPheAlaLeuTyrIleAspArgProValIleAsnGlyValIlePheSerThrValHis 60
DB 121 AGTAGATTGCTCTTACAAAGCGCCCAAGGCAATGGTGTCMAACCCAGCACCTGCAT 180
QY 61 ValIleSerThrProGlnAlaSerIleValIleSerCysValGlyGlnHisSerIleSer 80
```

Db 181 GGAATATCCAGCCCAAGGATCAAGGCTATAGCTGCAAGAGTCAACACATATATCC 240
QY 81 TTTThLeuSerArgAnGlnThrValValAlGluTyrThrHisAspLysAspThrAsp 100
Db 241 TACACTTTGTCAAGAAATCAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 101 MetPheGlnValAlGlyArgSerThrGluSerProLysAspPheValAlThrAspThrIle 120
Db 301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGACTCTGTTGTCAACAGACAGATT 360
QY 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 361 TCTGGCAGCCAGAAACAGAGCAGAAAGCCAGATCAACAGAGCAGATTCAGAGTTCCGC 420
QY 141 CysArgIleValCysAspArgAnGlnProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 421 TGCAGAGATGCTGTGCGACAGAAATGAACCTTACACAGCAGATATTCGCGCGGATTT 480
QY 161 AspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTrpLysAsnProAspGly 180
Db 481 GACTCTTCCAAAACATATTTCTTGAGTAAAGCAGAAAGTGGAACCCCGACGCGC 540
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db 541 CACATGATGGGCTCACTAATAGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 201 GluGluSerGlnProGlyValAlTPrArgLysIleSerValCysGlyAspValTyrThrLeu 220
Db 601 GAGGATCCCAAGCCCGGGGCTGTGGCGAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGlnThrAsnValLeu 240
Db 661 CGAGAAACCGAGTCCGCGCCAGCAACAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGT 720
QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuThrArgThrAlaAspGly 260
Db 721 CAGAGCGGCTCTCTCATGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
Db 781 CTTTTCATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 840
QY 281 ArgProGlnCysArgProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
Db 841 CGGCTCAAGTGTCTGTGGGGCTCAACACCTGCGCTTCCCGACATCAACAGAAAGAG 960
QY 301 ValAlaGluGluLysGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
Db 901 GTGGTGAAG 960
QY 321 AsnTrpGlyHisArgSerAspThrGlnAlaAsnGlnLysArgLysCysProMetCysArgThr 340
Db 961 AACTGGGGGCGATCGAGTGAACAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 341 ValGlyProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAlaAlaGly 360
Db 1021 GTGGGCGCCATATGCTCTCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlnLysSerAlaLysTyr 380
Db 1081 CCGCCAACTCATCTTCACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db 1141 TGGTCTCAAGTCCCGTTGCTCATGAACTCAAGCATTTACCGTGTGTGTGTGTGTGTGT 1200
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
Db 1201 GCTACACAGCTGGTGGGAGCAAACTGCATCAAAATTAATTTCCAAAGGTCCAAATGAC 1260

RESULT 8
US-09-843-905A-5

Sequence 5, Application US/09843905A
Patent No. US200201683A1
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Cosman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843,905A
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,198
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1260
TYPE: DNA
ORGANISM: Mus musculus
US-09-843-905A-5
Alignment Scores:
Pred. No.: 7,84e-246 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.38 Conservative: 10
Best Local Similarity: 95.004 Mismatches: 10
Query Match: 95.004 Indels: 1
Gaps: 1
DB: 9
US-10-041-030-4 (1-420) x US-09-843-905A-5 (1-1260)
QY 1 MetPheSerProGlyGlnGluHisCysAlaProAsnLysGluProValLysTyrGly 20
Db 1 ATGTTTCCCGGGCCAGAGAGAAACCAGCGCCCAACAGAGAGCGGGAATACGGG 60
QY 21 GluLeuValAlaLeuGlyTyrAnGlyAlaLeuProAnGlyAspArgLysArgLys 40
Db 61 GAGCTGTGTCTGTGGGATCAATGTCTTAACTGAAGTGTGTGTGTGTGTGTGTGTGTGT 120
QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db 121 AGCAGATTGGCTCTATAGCGGACCTACGCGCAGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
Db 181 ATGCTTCCACACAG 240
QY 81 TTTThLeuSerArgAnGlnThrValValAlGluTyrThrHisAspLysAspThrAsp 100
Db 241 TACACTTTGTCAAGAAATCAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 101 MetPheGlnValAlGlyArgSerThrGluSerProLysAspPheValAlThrAspThrIle 120
Db 301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGACTCTGTTGTCAACAGACAGATT 360
QY 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 361 TCTGGCAGCCAGAAACAGAGCAGAAAGCCAGATCAACAGAGCAGATTCAGAGTTCCGC 420
QY 141 CysArgIleValCysAspArgAnGlnProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 421 TGCAGAGATGCTGTGCGACAGAAATGAACCTTACACAGCAGATATTCGCGCGGATTT 480
QY 161 AspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTrpLysAsnProAspGly 180
Db 481 GACTCTTCCAAAACATATTTCTTGAGTAAAGCAGAAAGTGGAACCCCGACGCGC 540
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db 541 CACATGATGGGCTCACTAATAGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 201 GluGluSerGlnProGlyValAlTPrArgLysIleSerValCysGlyAspValTyrThrLeu 220
Db 601 GAGGATCCCAAGCCCGGGGCTGTGGCGAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGlnThrAsnValLeu 240
Db 661 CGAGAAACCGAGTCCGCGCCAGCAACAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGT 720
QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuThrArgThrAlaAspGly 260
Db 721 CAGAGCGGCTCTCTCATGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
Db 781 CTTTTCATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 840
QY 281 ArgProGlnCysArgProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
Db 841 CGGCTCAAGTGTCTGTGGGGCTCAACACCTGCGCTTCCCGACATCAACAGAAAGAG 960
QY 301 ValAlaGluGluLysGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
Db 901 GTGGTGAAG 960
QY 321 AsnTrpGlyHisArgSerAspThrGlnAlaAsnGlnLysArgLysCysProMetCysArgThr 340
Db 961 AACTGGGGGCGATCGAGTGAACAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 341 ValGlyProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAlaAlaGly 360
Db 1021 GTGGGCGCCATATGCTCTCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlnLysSerAlaLysTyr 380
Db 1081 CCGCCAACTCATCTTCACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db 1141 TGGTCTCAAGTCCCGTTGCTCATGAACTCAAGCATTTACCGTGTGTGTGTGTGTGTGT 1200
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
Db 1201 GCTACACAGCTGGTGGGAGCAAACTGCATCAAAATTAATTTCCAAAGGTCCAAATGAC 1260
QY 420 598 GAGGATCCCAAGCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 657

```

QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGlyIleuSerGluThrAsnValLeu 240
DB 658 CGAGAACACAGAGTCCGCGCCACGACGAGAGGGAAGCTGTGGAAAGTGAACCAACGTCCTG 717
QY 241 GlnAspGlySerLeuIleAspLeuCyseGlyValIleThrLeuLeuTrpArgThrAlaAspGly 260
DB 718 CAAGACGGCTCCCTCATGTGACCTGTGTGGGGCCACTCTCTCGAGAAACCGCAGATGGC 777
QY 261 LeuPheIleThrProThrGlnLysHsiIleGluAlaLeuArgGlnGlnIleLeuAlaIle 280
DB 778 CTTTTCACGCTCTCTCAAGAACATAGAACCCCTCGGACGAGATCAATGCAAGCC 837
QY 281 ArgProGlnCyseProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
DB 838 CGACCCCAAGTGCCTCGGGGCTTTAACCTCGCTTCCACGATCAACCGGAGGA 897
QY 301 ValValGlnGluLysGlnProTrpAlaIleLysSerCyseGlyValHsiIleGlyTrpHis 320
DB 898 GTGGTGAAGAGAGAGCCCTGGGCACTGAGCTGGCGGCATGTGACCGCTAC 957
QY 321 AsnTrpGlyHsiAspSerAspThrGluAlaAsnGluArgGluLysProMetCysArgThr 340
DB 958 AGCTGGGGCCATCGAGCGACCGGAGCCAGACGAGAGGAGTGTCCATGTGCAGACT 1017
QY 341 ValGlyProTrpValProLeuTrpLeuGlyCyseGluAlaGlyPheTrpValAspAlaGly 360
DB 1018 GTGGGCCCCCTACGCTCTCTGTGGCTGTGGCTGAGGAGATTTATGTGATGCGGA 1077
QY 361 ProProThrHsiAlaPheThrProCyseGlyValIleValCysSerGluLysSerAlaLysTrp 380
DB 1078 CCCCCAATCAACGCTTTCACCCCTCGGGCACGTGTTCAGAAAAGCTCGCAAGTAC 1137
QY 381 TrpSerGlnIleProLeuProHsiIleGlyThrHsiAlaPheHsiAlaIleCysProPheCys 400
DB 1138 TGGTCGACATCCCATCTGCCCGACGAGACGACGCTTCATGCGCTGTCCGTCTGC 1197
QY 401 AlaThrGlnLeuValGlyGlnAsnCyseIleLysLeuIlePheGlnGlyProIleAsp 420
DB 1198 GCCACGACCTGGTGTGTGAACGAACTCATTAATTTTCCAAAGTCCAGTGAC 1257

```

```

RESULT 9
US-10-317-250-5
; Sequence 5, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; APPLICANT: Li, Xiaoxia
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; CURRENT FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-317-250-5

```

```

Alignment Scores:
Pred. No.: 7, 84e-246 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.38 Conservative: 10
Best Local Similarity: 95.008 Mismatches: 10
Query Match: 95.008 Indels: 1
DB: 15 Gaps: 1

```

```

US-10-041-030-4 (1-420) x US-10-317-250-5 (1-1260)

```

```

QY 1 MetPheSerProGlyGlnGlnGluHsiCysAlaProAsnLysGluProValLysTrpGly 20
DB 1 ATGTTTTCCTCCGCGCCAGAGAAACCAACGCCCCCAACAGAGAGCCGGTGAATACGGG 60

```

```

QY 21 GlnLeuValValLeuGlyTrpAsnGlyValAlaLeuProAsnGlyAspArgGlyValArgLys 40
DB 61 GAGCTGGTGTCTCTGGGTTACATGTGCTTTACTTAATGTGACAGGGGCGAGAGAA 120
QY 41 SerArgPheAlaLeuTrpLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
DB 121 AGCAGATTGGCTCTATPAGCGGACCTTACGAGTGGTGTCAACCCGACCAATCCAC 180
QY 61 ValIleSerThrProGlnAlaSerLysValIleSerCyseGlyValHsiIleSerIleSer 80
DB 181 ATGGTCTCACACACAGCGCTCCAGGCGCACTCCAGAGGACATCACAGCATATCG 240
QY 81 TrpThrLysSerArgAsnGlnThrValValGlyTrpHsiAspLysAspThrAsp 100
DB 241 TACAGCTTGTACGAGACCGACGCTAGTGTGAGTACACACGATTAACACGAGAC 300
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
DB 301 ATGTTTCAGGTGGCAGGTCAACAGAAAGCCCATTTGACTTGTGTGCACACACGGTT 360
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 361 TCCGCGCGTCAGAAAGAAAGT--GCCAGATCACACAGACCATCTTAAAGTTTCGA 417
QY 141 CysArgIleValCysAspArgAsnGluProTrpTrpAlaArgIlePheAlaIleGlyPhe 160
DB 418 TCCAGGATCGTGTGACAGAACGACCATATACGACACGATTTTGGCGGACAGATTG 477
QY 161 AspSerSerLysAsnIlePheLeuGlyGlyLysValAlaLysTrpLysAsnProAspGly 180
DB 478 GATTCCTCCAAAATAATCTTTCTTGAGAGAAAGCAGCAAAATGAAAAACCTGATGA 537
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
DB 538 CACATGATGACCTACATCAACATGTGTCTTGAATGACACCCGCAAGAGAGCTTCAC 597
QY 201 GlnGluSerGlnProGlyValTrpArgGluIleSerValCysGlyAspValTrpThrLeu 220
DB 598 GAGGAATCCAGCTCGAGTCTGAGAGAGATCTGTCTGTGGGAGATGTGTACACTTG 657
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGlyIleuSerGluThrAsnValLeu 240
DB 658 CGAGAGACCAAGGTCCGCGCCACGACGAGGGAAGCTGTGGAAGTGAACCAACGTCCTG 717
QY 241 GlnAspGlySerLeuIleAspLeuCyseGlyValIleThrLeuLeuTrpArgThrAlaAspGly 260
DB 718 CAAGACGGCTCCCTCATGTGACCTGTGTGGGGCCACTCTCTGAGAAACCGCAGATGGC 777
QY 261 LeuPheIleThrProThrGlnLysHsiIleGluAlaLeuArgGlnGlnIleLeuAlaIle 280
DB 778 CTTTTCACGCTCTCTCAAGAACATAGAACCCCTCGGACGAGATCAATGCAAGCC 837
QY 281 ArgProGlnCyseProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
DB 838 CGACCCCAAGTGCCTCGGGGCTTTAACCTCGCTTCCACGATCAACCGGAGGA 897
QY 301 ValValGlnGluLysGlnProTrpAlaIleLysSerCyseGlyValHsiIleGlyTrpHis 320
DB 898 GTGGTGAAGAGAGAGCCCTGGGCACTGAGCTGGCGGCATGTGACCGCTAC 957
QY 321 AsnTrpGlyHsiAspSerAspThrGluAlaAsnGluArgGluLysProMetCysArgThr 340
DB 958 AGCTGGGGCCATCGAGCGACCGGAGCCAGACGAGAGGAGTGTCCATGTGCAGACT 1017
QY 341 ValGlyProTrpValProLeuTrpLeuGlyCyseGluAlaGlyPheTrpValAspAlaGly 360
DB 1018 GTGGGCCCCCTACGCTCTCTGTGGCTGTGGCTGAGGAGATTTATGTGATGCGGA 1077
QY 361 ProProThrHsiAlaPheThrProCyseGlyValIleValCysSerGluLysSerAlaLysTrp 380
DB 1078 CCCCCAATCAACGCTTTCACCCCTCGGGCACGTGTTCAGAAAAGCTGCAAGTAC 1137

```

```

QY 381 TrpSerGlnIleProleuProHisGlyThrHisAlaPheHisAlaIaCysProPheCys 400
DB 1138 TGGTCCAGATCCCACTGCCCCACGAGACGACGCTTCATCCGCTGCTCGGTTCTGC 1197
QY 401 AlaThrGlnLeuValGlyGlnGlnAnCysIleLeuLeuIlePheGlnGlyProIleAsp 420
DB 1198 GCCAGCGACGCTGTTGTGTGAACAGAACTGCATCAATTAATTTTCCAGAGTCCAGTGGAC 1257

RESULT 10
US-10-085-117-285
; Sequence 285, Application US/10085117
; Publication No. US200302334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 52945200121
; CURRENT APPLICATION NUMBER: US/10/085,117
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 285
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-085-117-285

Alignment Scores:
Pred. No.: 7,846-246 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.38% Conservative: 10
Best Local Similarity: 95.00% Mismatches: 1
Query Match: 15 Gaps: 1

US-10-041-030-4 (1-420) * US-10-085-117-285 (1-1260)
QY 1 MetPheSerProGlyGlnGlnGlnHisCysAlaProAsnIleGluProValIleTyRgIy 20
DB 1 ATGTTTCCCGGCGGACGAGGAGAACCCAGCGCCCAACAGAGCGCGTGAATACCGG 60
QY 21 GluLeuValIleLeuGlyTyRangIyAlaLeuProAsnIleAspArgGlyArgArgIy 40
DB 61 GAGCTGTGTCTCTGGGCTACAAATGCTTACTTAATGCTGACAGGCGGACGAGGAA 120
QY 41 SerArgPheAlaLeuTyRlyAspArgProIleAlaAsnGlyValIleProSerThrValHis 60
DB 121 AGCAGATTGCGCTCTAATAGCGGACCTACGCCAGTGTGTCAAAACCCACGACAAATCCAC 180
QY 61 ValIleSerThrProGlnAlaSerIleValIleSerCysLeuGlyGlnHisSerIleSer 80
DB 181 ATGGTTCACACCAACAGCGGCTCCAGCCATCAGCTCCAGAGACATCAGCATATCG 240
QY 81 TyrThrLeuSerArgAngInThrValIleValIleGlyTyRThrHisAspIleAspThrAsp 100
DB 241 TACACCTGTCTACGAGCGACAGACGTAAGTGTGAGTACACACATATAAGACACAGGAC 300
QY 101 MetPheGlnValIleArgSerThrGlnSerProIleAspPheValIleThrAspThrIle 120
DB 301 ATGTTTCAGGTGGCGGCTCAACAGAAAGCCCATGACTCTGTGTCAACAGACGAGTT 360
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 361 TCCGGCGGTCAAGAACGAGT---GCCAGATCACAACAGGACCATCTCTAGTTCCGA 417
QY 141 CysArgIleValIleCysAspArgAngIleProTyRThrAlaArgIlePheAlaIleGlyPhe 160
DB 418 TCCAGAGATGCTGTGACAGGAAAGGACATATACAGCAGCATATATCCGCGGACGATTC 477
QY 161 AspSerSerIleAsnIlePheLeuGlyGlyValAlaAlaIleTyRTrpIleAsnProAspGly 180

```

```

DB 478 GATCTTCCAAAATATCTTTCTGAGAGAAAGACAAATGAAAACCTGTATGA 537
QY 181 HisMetAspGlyLeuThrThrAngIyValIleuValIleuHisProArgIyGlyPheThr 200
DB 538 CACATGATGAGACTCACTACCAATGGGTCTCTAGTATGACACCCGACAGAGGCTTCACC 597
QY 201 GluGlnSerGlnProGlyValIleTrpArgIleIleSerValCysGlyAspValTyRThrIle 220
DB 598 GAGAAATCCACGCTGAGTGTGAGAGAGATCTGTGTGTGGGATGTATACACTTG 657
QY 221 ArgGluThrArgSerAlaGlnGlnIleArgGlyIleLeuValIleSerGlnThrAsnValIle 240
DB 658 CAGAGACCAAGTTCGCGCCACGACGAGGAGAAAGCTGTGTGAAGTACAGCAAGCTCTG 717
QY 241 GlnAspGlySerIleuIleAspLeuCysGlyAlaIleThrLeuLeuTrpArgThrIleAspGly 260
DB 718 CAAGACGCTCTCTCATTAAGCTGTGTGGGACACTCTCTGTGAGAACCCGACAGATGAC 777
QY 261 LeuPheIleThrProThrGlnIleHisIleGluAlaLeuArgGlnIleIleAsnAlaIle 280
DB 778 CTTTTCACGCTCTTACTGAGAGACATGAAAGCCCTCCGACGAGATCAATGACGCC 837
QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgIyGln 300
DB 838 CACCCCAATGCGCCGTGGGCTTAAACCTTGCCTTCCACGATCAACCGAGAGAA 897
QY 301 ValIleGlnGlyIleProIleProIleTyRLeuSerCysGlyIleValIleGlyTyRHis 320
DB 898 GTGTGTGAAGAGACAGGACCTGGGCTGAGTCTGAGCTGGGACATGTGACGCGCTACAC 957
QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnIleuArgIleCysProMetCysArgThr 340
DB 958 AGCTGGGCGATCGAGCGACCGGAGGCAACGAGAGGAGTGTCCATGTGCAGAGCT 1017
QY 341 ValIleProTyRValProleuTrpLeuGlyCysGluAlaGlyPheTyRValAspAlaGly 360
DB 1018 GTGGGCGCTTACGCTCTCTGTGTGGCTGTGAGCGAGATTTATATCATGCGGGA 1077
QY 361 ProThrHisAlaPheThrProCysGlyIleValCysSerGluIleSerAlaIleTyR 380
DB 1078 CCCCACATCAGCTTTCACCCCTGCGGCGACGCTGTGTAAGAAAGCTGCGCAAGTAC 1137
QY 381 TrpSerGlnIleProleuProHisGlyThrHisAlaPheHisAlaIaCysProPheCys 400
DB 1138 TGGTCCAGATCCCACTGCCCCACGAGACGACGCTTCATCCGCTGCTCGGTTCTGC 1197
QY 401 AlaThrGlnLeuValGlyGlnGlnAnCysIleLeuLeuIlePheGlnGlyProIleAsp 420
DB 1198 GCCAGCGACGCTGTTGTGTGAACAGAACTGCATCAATTAATTTTCCAGAGTCCAGTGGAC 1257

RESULT 11
US-10-258-703-5
; Sequence 5, Application US/10258703
; Publication No. US20040034199A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-US
; CURRENT APPLICATION NUMBER: US/10/258,703
; PRIOR FILING DATE: 2002-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-258-703-5

Alignment Scores:
Pred. No.: 7,846-246 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.38% Conservative: 10

```

Best Local Similarity: 95.00% Mismatches: 10
Query Match: 95.00% Indels: 1
DB: 16 Gaps: 1
US-10-041-030-4 (1-420) x US-10-258-703-5 (1-1260)

QY 1 MetPheSerProGlyGlnGluGluHisCybA1AProAenlySGlUProValIlySTyGly 20
DB 1 ATGTTTTCCTCCGGGCGCAGAGAGAACCCAGCGCCCAACAGAGAGCGGTGAATATCGGG 60
QY 21 GluLeuValIalleuGlyTyTyAsnGlyAlaLeuProAsnGlyAspArgGlyValArgGly 40
DB 61 GAGCTGGTGCTCTGGGGGTACATAGTGCTTAAATGTGTACAGGGGGCGAGGAA 120
QY 41 SerArgPheAlaLeuTyTyAsnArgProlySAlaAsnGlyValIlyProSerThrValHis 60
DB 121 AGCAGATTGGCCCTCATAGGCGGACCTAGCCAGTGGTGTCAAAACCCAGCAATCCAC 180
QY 61 ValIleSerThrProGlnAlaSerIlySAlaIleSerCybLySGlyGlnHisSerIleSer 80
DB 181 ATGCTTCACACACACAGCGCGTCCAGGCATAGCTCCAGAGAGCATCACAGCATATCG 240
QY 81 TyTrhLeuSerArgAsnGlnThrValIalleuGlyTyTyTrhIleAspIlyAspThrAsp 100
DB 241 TACACGTTGTCAACGAGCCAGACGGTAGTGGTGAGTACACACCATTAAGACACGGAC 300
QY 101 MetPheGlnValIalleuSerThrGluSerProIleAspPheValIalleuPheThrIle 120
DB 301 ATGTTTCAGGTGGCGAGGTCAACAGAAACCCCATGATCTTGAGTCAACAGACGGTT 360
QY 121 SerGlySerGlnAsnThrAspGluIalleuGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 361 TCCGGCGGTCAAGACGAAAT---GCCAGATCACACAGACCATCTCTAGGTTCCGA 417
QY 141 CybAglIleValIalleuAspArgAsnGluProTyTyTrhAlaArgIlePheAlaIleGlyPhe 160
DB 418 TCCAGAGTGTGTGTACAGAGACGAGCCATATACAGCAATATTCCTGGCAGAGATTCC 477
QY 161 AspSerSerIlyAsnIlePheLeuGlyGlyValIalleuValIalleuTyTrhIlyAsnProAspGly 180
DB 478 GATTCTTCCAAATAATCTTCTTGGAGAGAAAGCAGAAATGGAAATCCCTGATGGA 537
QY 181 HisMetAspGlyLeuThrIleAsnGlyValIalleuValMetHisProArgGlyGlyPheThr 200
DB 538 CACATGGATGGAATCACTACATGATGTCCTATGATGATCAACCGGAAAGAGGCTTCAAC 597
QY 201 GluGluSerGlnProGlyValIalleuArgGluIleSerValIalleuAspValIlyTyTrhLeu 220
DB 598 GAGGAAATCCACCTGGAGTCTGGAGAGAGATCTGTCTGTGGGATGTGTACACCTTG 657
QY 221 ArgGluThrArgSerAlaGlnIleArgGlyValIalleuValIalleuGluSerGluThrAsnValIleu 240
DB 658 CCAAGAGACAGATCCGCGCCAGACAGAGGAGGAAAGCTGGTGAAGATGAGACCAAGCTCG 717
QY 241 GlnAspGlySerLeuIleAspLeuCybGlyAlaIleThrLeuIleTrpArgThrAlaAspGly 260
DB 718 CAAAGAGGCTCCCTATGATCCGTGTGGGCAACTCTCTGTGAGAAACCCAGATGCG 777
QY 261 LeuPheHisThrProThrGlnIlyHisIleGluAlaLeuArgGlnIlyLeuAsnAlaIle 280
DB 778 CTTTTCACGCTCTACTAGAGACCATAGAGAGCCCTCCGACAGATCAATGACAGCC 837
QY 281 ArgProGlnCybProValIalleuAsnThrLeuAlaPheProSerIleAsnArgIlyGlu 300
DB 838 GCAACCCCAATGCGCCGTGGGCTTAAACACCTGGCTTCCAGCATCAACCGAGAGGA 897
QY 301 ValIalleuGluIlySGlnProTrpAlaTyTyLeuSerCybGlyHisValHisGlyTyTrhHis 320
DB 898 GTGGTGAAGAGAGAGAGCGCTGGGCAATCTGAGCTGGGCGCATGTGACAGGCTACAC 957
QY 321 AsnTrpGlyHisIleAspSerAspThrGluAlaAsnGluArgGluCybProMetCybArgThr 340
DB 958 ACCTGGGCGCATCGAGCGAGCGGAGCAACAGAGAGAGAGTGTCTCCATGTGACGAGACT 1017

QY 341 ValGlyProTyTyValIalleuThrLeuGlyGlyGluAlaIlePheTyTyValIalleuGly 360
DB 1018 GTGGGCGCTTACGCTCTCTGAGCTGTGTAGGAGAGATTTATGTCATCGGGGA 1077
QY 361 ProProThrHisAlaPheThrProCybGlyHisValIalleuSerGluIlySerAlaIlyTyTr 380
DB 1078 CCCCCAATCAACGCTTTCACCCCTCGGGGACGCTGTTCAGAAAGTCTGCCAAGTAC 1137
QY 381 TrpSerGlnIleProLeuProHisGlyTyTrhHisAlaPheHisAlaIalleuCybProPheCyb 400
DB 1138 TGGTGCAGATCCCATCTGCCCAAGAGACGAGGTTTCATGCGCTGTCTGTC 1197
QY 401 AlaThrGlnLeuValIalleuGlnIleAsnCybIleIlySAlaIlePheGlnGlyProIleAsp 420
DB 1198 GCCACGACGCTGTGTGTGAAACAGACTGCATCAATGATTTCCAAAGTCCATGAGAC 1257

RESULT 12
US-10-197-666A-133
; Sequence 133, Application US/1019766A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAH KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: Elk1 phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 133
; LENGTH: 1717
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (218)..(1474)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1288..1301
; OTHER INFORMATION: n-A or C or G or T
US-10-197-666A-133

Alignment Scores:
Pred. No.: 4,25e-238 Length: 1717
Score: 2111.50 Matches: 388
Percent Similarity: 95.24% Conservative: 12
Best Local Similarity: 92.36% Mismatches: 19
Query Match: 92.21% Indels: 1
DB: 14 Gaps: 1

US-10-041-030-4 (1-420) x US-10-197-666A-133 (1-1717)

QY 1 MetPheSerProGlyGlnGluGluHisCybA1AProAenlySGlUProValIlySTyGly 20
DB 1 ATGTTTTCCTCCGGGCGCAGAGAGAACCCAGCGCCCAACAGAGAGCGGTGAATATCGGG 277
QY 21 GluLeuValIalleuGlyTyTyAsnGlyAlaLeuProAsnGlyAspArgGlyValArgGly 40
DB 278 GAGCTGGTGCTCTGGGGGTACATAGTGCTTAAATGTGTACAGGGGGCGAGGAA 337
QY 41 SerArgPheAlaLeuTyTyAsnArgProlySAlaAsnGlyValIlyProSerThrValHis 60


```

Db      338 AGCAGATTACCTCTATTAAGCGACTACGACGAGTGTGTCAAAACCCAGCAATATCCAC 397
Qy      61 ValIleSerThrProGlnAlaSerIlyValIleSerCysValGlyGlnHisSerIleSer 80
Db      398 ATGGTCTCCACACCAAGGCGTCCAGGCGCATCAGCTCCAGAGACATCAACGATATGCG 457
Qy      81 TyrThrLeuSerArgAsnGlnThrValValValGluTyrThrHisAspIlyAspThrAsp 100
Db      458 TACACGCTTTCACGAGCGACGAGAGTGTGTGAGTATACACACATTAAGCACAGAC 517
Qy      101 MetPheGlnValAlaYarSerThrGlnSerProIleAspPheValValThrAspThrIle 120
Db      518 ATGTTTCAGGTGGGCGAGTCAACAGAAAGCCCATTTGACTCTGTGTCACAGACGGTT 577
Qy      121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      578 TCCGGCGGTCAAGAACAGAT--GCCAGATCACACAGACCAATCTCTAGTTCCGA 634
Qy      141 CysAlaGlyLeuAlaCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db      635 TCGAGGATCGTGTGACAGAGACGAGCATACAGACGACGATATTCGGCGCAGGATTC 694
Qy      161 AspSerSerIlyAsnIlePheLeuGlyGlyValAlaAlaIleTyrPheAspProAspGly 180
Db      695 GATTCTTCAAAAATATCTTTCTTGGAGAGAAAGCAGCAAAATGGAAAAACCTGATGGA 754
Qy      181 HisMetAspGlyLeuThrThrAsnGlyValIleValMetHisAspArgGlyGlyPheThr 200
Db      755 CACATGATGAGACTCACTACCAACGGTGTCTGTATGACCCGCAAGAGAGCTTCAC 814
Qy      201 GluGlnSerGlnProGlyValIleTyrAspGluIleSerValCysGlyAspValIleThrLeu 220
Db      815 GAGGAATCCACAGCTGAGTGTGAGAGAGATCTCTGTGTGGGATGATACACCTTG 874
Qy      221 ArgGluThrArgSerAlaGlnGlnIleArgIlyLeuValGlnSerGlnThrAsnValIleu 240
Db      875 CGAGAACAGAGTGGCCCGACAGAGAGAAACCTGTGAGAAAGTAAAGCAACGCTCG 934
Qy      241 GlnAspGlySerLeuIleAspLeuGlyValAlaThrLeuLeuTyrArgThrAlaAspGly 260
Db      935 CAAGACGGCTCCCTCATTAACCTGTGTGGGCGCACCTCTCTGAGAACCCGACAGTGG 994
Qy      261 LeuPheHisThrProThrGlnIlyHisIleGlnAlaLeuArgGlnIleAsnAlaAla 280
Db      995 CTTTTTCACGCTCTCTACTCAGAAAGCACAATAGAACCTCCGCGCAGAGATCAATGACGCC 1054
Qy      281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgIlyGlu 300
Db      1055 CAGCCCAAGTGGCCCTGTGGGCTTAAACCTGTGGCTTCCCGACATCAACCGGAAGAA 1114
Qy      301 ValIleGluGlyLeuGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
Db      1115 GTGGTGAAGAGAGAGAGCGCTGTGGCATACCTAGCTGTGGCGCATGTGCACGGCTACAC 1174
Qy      321 AsnThrGlyHisArgSerAspThrGluAlaAsnGlnArgGluCysProMetCysArgThr 340
Db      1175 CACTGGGGCCATCGAGCAGCAGAGAACCAAGAGAGAGTGTCCCATGTGCATGACT 1234
Qy      341 ValIleProTyrValIleProLeuTyrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
Db      1235 GTGGGCGCTACCTCTCTGTGGCTGTGAGCGAGATTTATGTCCGANNNNNN 1294
Qy      361 ProProThrHisIleAspThrProCysGlyHisValCysSerGlnIlySerAlaIlyTyr 380
Db      1295 NNNNNNNCATGTTTTCACCCCTCGGGCAGTGTCTTCAGAGAGCTGTGCAAGTTC 1354
Qy      381 TrpSerGlnIleProLeuProHisGlyThrHisIlePheHisAlaAlaCysProPheCys 400
Db      1355 TGGTCCAGATCCCATGCGCCACGGAACGACGCGCTTTCACCCCGCTGTCCGTTCTGC 1414
Qy      401 AlaThrGlnLeuValAlaGlyGluGlnAsnGlyIleTyrLeuIlePheGlnGlyProIleAsp 420

```

```

Db      1415 GCCACGACGTGTGTGTAACAGAACTGCATCAATTGATTTTCCAGGTCCAGTGCAC 1474
RESULT 13
US-10-085-117-284
; Sequence 284, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Moritz, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
; NAME/KEY: Variation
; FEATURE:
; LOCATION: (1)...(1734)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-284

```

Alignment Scores:

Pred. No.:	4,31e-238	Length:	1734
Score:	2111.50	Matches:	388
Percent Similarity:	95.24%	Conservative:	12
Best Local Similarity:	92.38%	Mismatches:	19
Query Match:	92.21%	Indels:	1
DB:	15	Gaps:	1

US-10-041-030-4 (1-420) x US-10-085-117-284 (1-1734)

```

Qy      1 MetPheSerProGlyGlnGluGlnHisCysValAspAsnIlyGluProValIlyTyrGly 20
Db      218 ATGTTTCCCGCGGCGCAGAGAGAACCCAGCGCCCAACAGAGCGGTGAATACAG 277
Qy      21 GluLeuValIleGlyTyrAsnGlyAlaLeuProAsnGlyAspArgIlyArgAlly 40
Db      278 GAGCTGTGTCTGTGGTGAATGTCCTTAACCTAAATGTGACAGGGGCAAGAGAA 337
Qy      41 SerArgPheAlaLeuTyrIlyAspProIlyAlaAsnGlyValIlyProSerThrValHis 60
Db      338 AGCAGATTACCTCTATTAAGCGACCTACGCAAGTGTGTAAACCCAGCAATATCCAC 397
Qy      61 ValIleSerThrProGlnAlaSerIlyValIleSerCysValGlyGlnHisSerIleSer 80
Db      398 ATGGTCTCCACACCAAGGCGTCCAAAGGCGCATCAGCTCCAGAGACATCAACGATATG 457
Qy      81 TyrThrLeuSerArgAsnGlnThrValValValGluTyrThrHisAspIlyAspThrAsp 100
Db      458 TACACGCTTTCACGAGCGACGAGAGTGTGTGAGTATACACACATTAAGCACAGAC 517
Qy      101 MetPheGlnValAlaYarSerThrGlnSerProIleAspPheValValThrAspThrIle 120
Db      518 ATGTTTCAGGTGGGCGAGTCAACAGAAAGCCCATTTGACTCTGTGTCACAGACGGTT 577
Qy      121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      578 TCCGGCGGTCAAGAACAGAT--GCCAGATCACACAGACCAATCTCTAGTTCCGA 634
Qy      141 CysAlaGlyLeuAlaCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db      635 TCGAGGATCGTGTGACAGAGACGAGCATACAGACGAGATATTCGGCGCAGGATTC 694
Qy      161 AspSerSerIlyAsnIlePheLeuGlyGlyValAlaAlaIleTyrPheAspProAspGly 180
Db      695 GATTCTTCAAAAATATCTTTCTTGGAGAGAAAGCAGCAAAATGGAAAAACCTGATGGA 754

```

```

QY 181 HisMeaSPGlyLeuThrThraNGlyValLeuValMeNI:ProarglyGlyPheThr 200
DB 755 CACATGGATGACTCACTACCAAGCTGCTCGTATGACACCCGAAAGAGCTTCACC 814
QY 201 GluGlySerGlnProGlyValItrArgGluIleSerValCyseGlyAspValItrThrLeu 220
DB 815 GAGGAATCCCACTGGAGTCTGGAGAGAGATCTGCTGTGGGATGTATAACCTTG 874
QY 221 ArgGluThrArgSerAlaGlnItrArgGlyValLeuValGlySerGluThrArgValLeu 240
DB 875 CGAAGAACCAAGTCCGCGCCAGCAAGAGGAGAGCTGTGTGAAGTGAAGCAACGCTTG 934
QY 241 GlnAspGlySerLeuIleAspLeuCyseGlyAlaThrLeuLeuItrArgThrAlaAspGly 260
DB 935 CAAGAAGCTCCCTCATTTGACCTGTGTGGGCCACTCTCTGGAGAACCCCAAGTGGC 994
QY 261 LeuPheNI:ThrProThrGlnIlyshIleGlyAlaLeuArgGlnIleAspAlaAla 280
DB 995 CTTTTCACGCTCTACTACAGAAAGCACATAGAAGCCCTCCGCGAGAGATCAATGACGC 1054
QY 281 ArgProGlnItrProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArglySerGlu 300
DB 1055 CAAACCCCAAGTCCCGTGGGCTTAAACACCTGGCTTCCCAAGCATCAACCGAAAGAA 1114
QY 301 ValValGluGluIlySerGlnProItrAlaIlyLeuSerCyseGlyIlyshValIshGlyIlyshIsh 320
DB 1115 GTGGTGAAGAAGAAAGAGCCCTGGGATACCTGAAGTGGGCCCATGTGACCGCTACCC 1174
QY 321 AsnItrPGLyIshArgSerAspThrGluAlaAsnGluArgGluCyseProMetCyseArgThr 340
DB 1175 CACTGGGGGCATCGAGGCAACACGGAAGCAACGAGAGGAGTGTCCCATGTGCAATGACT 1234
QY 341 ValGlyProItrValProLeuItrPheGluIlyCyseGluAlaGlyPheThrValAspAlaGly 360
DB 1235 GTGGGCCCCCTAGCTCCCTCTGTGGCTGTGAGCGAGATTATATCGAANNNNNN 1294
QY 361 ProItrThrAlaPheThrProCyseGlyIlyshValCyseSerGluIlySerAlaIlyshIlyr 380
DB 1295 NNNNNNNCTCATGTTTTCACCCCTGGCGGCAAGTGTTCAGAGAAGCTCCCAAGTAC 1354
QY 381 TrpSerGlnIleProLeuProNI:sglyThrAlaPheNI:AlaAlaCyseProPheCyse 400
DB 1355 TGTGTGCAGATCCCACTGCCCAAGCAAGCAAGCGCTTTCACGCGCTGTCCGTTCTGC 1414
QY 401 AlaThrGlnLeuValGlyGlnIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlysh 420
DB 1415 GCCAGCAGCTGTGTGTGAAGCAAGATCATCAATTAATTTCCAAAGTCCAGTGCAC 1474

```

RESULT 14
US-09-843-905A-3
; Sequence 3. Application US/09843905A
; Patent No. US2002016863A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 3
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-843-905A-3

Alignment Scores:
Pred. No.: 2,236-215 Length: 1257
Score: 1917.00 Matches: 342
Percent Similarity: 90.48% Conservative: 38

```

Best Local Similarity: 81.43% Mismatches: 38
Query Match: 83.71% Indels: 2
DB: 9 Gaps: 1
US-10-041-030-4 (1-420) x US-09-843-905A-3 (1-1257)
QY 1 MetPheSerProGlyGlnGluIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlysh 20
DB 1 ATGTTTCTCCGATCAAGAAATCAT-----CCATCTAAAGCAACGATAAATATGTT 54
QY 21 GluLeuValValLeuGlyIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlysh 40
DB 55 GAACCATGTGCTTGTGAATATGATGATCTCCCAACCGGTGAATGAGAAAGGAGAA 114
QY 41 SerArgPheAlaLeuIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlysh 60
DB 115 AGTAGTTTGTCTTGTAAAGACTTAAGGCAATAGGAGGAGAACCCAGCATGTGCT 174
QY 61 ValIleSerThrProGlnAlaSerIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlysh 80
DB 175 ATTGCTTGATCTCCCTCAGGCTGCAAGGCAATAGCAACAAAGACCATATGATATCA 234
QY 81 TyrThrIleSerArgAsnGlnThrValValValGluIlyshIlyshIlyshIlyshIlysh 100
DB 235 TATACTTATATCGGGCCAGACTGTGGTGTGAATATGATCAAGACAGCAACCGAT 294
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValItrAspThrIle 120
DB 295 ATGTTTCAGATTGGCGGTGCACTGAAGAGCCCATGATTTGTGTAAGTGAACGAGTT 354
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIlyshIlyshIlyshIlyshIlyshIlysh 140
DB 355 CCGTGAATCAAGAAATTCGATACAGACAGTACAGAAAGACATATCAAGATTGCC 414
QY 141 CyseArgIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlysh 160
DB 415 TGCAGAAATCAATATGAAAGCAAGAAATCTCCCTTACAGACAGCAAGTTTATGTCAGAGGTTT 474
QY 161 AspSerIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlyshIlysh 180
DB 475 GACTATCAAAAGAAATCTTCTTGGAGAAAGGCTCCCAATGGAAGACATCAGATGGA 534
QY 181 HisMeaSPGlyLeuThrThraNGlyValLeuValMeNI:ProarglyGlyPheThr 200
DB 535 CAGATGAGATGGTTGACACATATATGTTCTTGTGATGATCAACGCAAGGGTTACA 594
QY 201 GluGlySerGlnProGlyValItrArgGluIleSerValCyseGlyAspValItrThrLeu 220
DB 595 GAAGACTCCAAAGCCCTGGAATATGAGAGAAATATCGTGTGTGGAAATGTATTAAGCCTA 654
QY 221 ArgGluThrArgSerAlaGlnItrArgGlyValLeuValGlySerGluThrArgValLeu 240
DB 655 CGTGAACCAAGATCCGCTCAGCAAGAGAGAAATGTGTGAATTTGAATCCATATAGTTA 714
QY 241 GlnAspGlySerLeuIleAspLeuCyseGlyAlaThrLeuLeuItrArgThrAlaAspGly 260
DB 715 CAAGATGCTGTGTTAATGACCTGTGTGCAACATTTATGCGATGCAAGAGGC 774
QY 261 LeuPheNI:ThrProThrGlnIlyshIleGlyAlaLeuArgGlnIleAspAlaAla 280
DB 775 CTTTCCCACTCCCTACCGGTGAAGCATTTAAGAGCTTTAAGACGAAATCAATCAACACA 834
QY 281 ArgProGlnItrProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArglySerGlu 300
DB 835 CGACTTCAGTCCCTGTAAGGTTCAACACCTAGCATTTCTTAGATATGAAGAGAAAGAC 894
QY 301 ValValGluGluIlySerGlnProItrAlaIlyLeuSerCyseGlyIlyshValIshGlyIlyshIsh 320
DB 895 GTTGTAGATGAAGAAACAAACATGGGTATATCAATCGGCAATGATCATGCGTATCAT 954
QY 321 AsnItrPGLyIshArgSerAspThrGluAlaAsnGluArgGluCyseProMetCyseArgThr 340
DB 955 AACTGGGGAACAAAGAAAGACGTGATGCAAGATGTGTGAATGTCTTATGTGTAGTGTCT 1014

```

```

QY 341 ValGlyProTyrValProLeuTTPLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
Db 1015 GTTGGGCCCTATATCTCTGTGGCTGTGATGTGAAGCTGGATTTATGTGGACGCCGCGC 1074
QY 361 ProPheThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
Db 1075 CTTCCAAACCAATCGCTTACCGCCGTGTGGCATGTGTTCAGAAAAGACAACATGCCCTAT 1134
QY 381 TTPSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaCysProPheCys 400
Db 1135 TGGTCCCAAGTCCCACTTCTCTCATGTGATCACTTTTCATGACGCTGTCTCTTTTGT 1194
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
Db 1195 GCACATCAGTTGGCTGTGTGAACAAGCTACATCAGATTATTTTTCAGAGCACTCTAGAC 1254

RESULT 15
US-10-317-250-3
; Sequence 3, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; APPLICANT: Li, Xiaoxia
; TITLE OF INVENTION: HUMAN PBLINO POLYPEPTIDES
; FILE REFERENCE: 2990-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-317-250-3

Alignment Scores:
Pred. No.: 2,23e-215 Length: 1257
Score: 1917.00 Matches: 342
Percent Similarity: 90.48% Conservative: 38
Best Local Similarity: 81.43% Mismatches: 38
Query Match: 83.71% Indels: 2
DB: 15 Gaps: 1

US-10-041-030-4 (1-420) x US-10-317-250-3 (1-1257)
QY 1 MetPheSerProGlyGlnGluGlnHisCysAlaProAsnLysGluProValLysTyrGly 20
Db 1 ATGTTTTCCTCGATCAAGAAATCAT-----CAATCTAAAGCACACAGTAAATATGAT 54
QY 21 GlnLeuValIleuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGlyLys 40
Db 55 GAACATATTGTCTTAGATATATATGATCTCTCCAAACGATGATGAGGAAGAGAGAAA 114
QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db 115 AGAAGCTTTCCTTTGTTTAAAGACCTAAGCAAAATGGGGTGAAGCCACGACTGTGCAT 174
QY 61 ValIleSerThrProGlnIleSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
Db 175 ATTGCTTGTACTCTCAAGGTGCAAGGCAATAGCAACAACCAACCACTACATATCA 234
QY 81 TyrThrLeuSerArgAsnGlnThrValValIleGlyTyrThrHisAspLysAspThrAsp 100
Db 235 TATACTTTATCTCGGGCCAGACTGTGTGTGTAATATATCTCATGACAGCAACAGCAT 294
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120
Db 295 ATGTTTCAAGATTGGCCGGTGCAGTGAAGCCCAATGATTGTTTGTAGTAACCTGACACGTT 354
QY 121 SerGlySerGlnSerThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140

```

```

Db 355 CTTGAAATCCAAAGTAAATTTCTGATACAGTCAAGTCAAAAGCACTATTCAGATTGGC 414
QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 415 TGCAGATCATATATGTAACGGAAATCTCCCTTTACAGCAACGGATTATGTCTGACAGGTT 474
QY 161 AspSerSerLysAsnIlePheLeuGlyGlyValAlaLysTyrLysAsnProAspGly 180
Db 475 GACTCATCAAAAACATCTTTCTTGGGGAAGAGCGCCAAATGGAAAGCATCATGATGGA 534
QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleuValMetHisProArgGlyGlyPheThr 200
Db 535 CAGATGATGGCTTGAACCAATATGTTCTTTGTATGATCATCCAGCAATGGGTTTCA 594
QY 201 GlnGluSerGlnProGlyValTyrArgGlnIleSerValCysGlyAspValTyrThrLeu 220
Db 595 GAAGACTCAAGGCTGGAAATATGGAAGAAATATCGGTGTGGGAATGATTTAGCCTA 654
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
Db 655 CTTGAAACCAAGATCGGCTCAGCAGAGAGAAAAATGCTGAAATTTGAAACCAATCAGTTA 714
QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrArgThrAlaAspGly 260
Db 715 CAAGATGGCTCTTAAATTAACCTCTGTGTGCAACATTTATTTGCGTACTGCAGAGGC 774
QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGlnIleAsnAlaAla 280
Db 775 CTTTCCCAACATCTCAACCGTGAAGCATTTAGAAGCTTTAAGACAGAAATCATGACAGA 834
QY 281 ArgProGlnCysProValAlaLysLeuThrLeuAlaPheProSerIleAsnArgLysGlu 300
Db 835 CGACCTCAGTGCCTGTAGGTTTCAACACATTCCTTAAGTATGAAGAGAAAGAC 894
QY 301 ValIleGluGlyLysGlnProTyrAlaTyrLysSerCysGlyHisValHisGlyTyrHis 320
Db 895 GTTGTAGATGAAGAAACACATGGGTATATCTTAACCTGGCCCATGTACATGGCTATCAT 954
QY 321 AsnTyrGlyHisArgSerAspThrGlnAlaAsnGlyArgGlyCysProMetCysArgThr 340
Db 955 AACTGGGGAACAACAAGAACGTGATGGCAAAAGATCGTAATGTCTTATGTGTAGTGT 1014
QY 341 ValGlyProTyrValProLeuTTPLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
Db 1015 GTTGGGCCCTATATCTCTGTGGCTGTGATGTGAAGCTGGATTTATGTGACGCCGCGC 1074
QY 361 ProPheThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
Db 1075 CTTCCAAACCAATCGCTTACCGCCGTGTGGCATGTGTTCAGAAAAGACAACATGCCCTAT 1134
QY 381 TTPSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaCysProPheCys 400
Db 1135 TGGTCCCAAGTCCCACTTCTCTCATGTGATCACTTTTCATGACGCTGTCTCTTTTGT 1194
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
Db 1195 GCACATCAGTTGGCTGTGTGAACAAGCTACATCAGATTATTTTTCAGAGCACTCTAGAC 1254

```

Search completed: December 10, 2004, 00:00:36
 Job time : 872 secs

This Page Blank (uspto)

CL

P.D. 03-09-2001
1-2 (2)

ID AAH34231 standard; cDNA; 864 BP.

XX
AC AAH34231; XP-002235785
XX

DT 03-SEP-2001 (first entry)

XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:1313.

XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.

XX
OS Homo sapiens.

XX
PN WO200122920-A2.

XX
PD 05-APR-2001.

XX
PP 28-SEP-2000; 2000WO-US26524.

XX
PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;

XX
DR WPI; 2001-235357/24.

DR P-PSDB; AAG74826.

XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers .

XX
PS Claim 1; Page 3049; 9803pp; English.

XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX
SQ Sequence 864 BP; 247 A; 142 C; 157 G; 314 T; 3 other;

ctgcttaatt gaagtgtaat atagcttcta gaactgttac ctgcagttct atggttttct	60
ttcacttctt ttctttttta aagccattct gttcttttga tgtgcttgaa aggggtgtgt	120
attacaccat tgttaatgct gggtaaaaaa taacttttct cagccttgcc tcataacagt	180
ggaattttctg atagacaaac cacaggaact tgaatttaag ccaaatccat ctccatccct	240
ttactgtcaa tcttctgtcc cagtacttta gcccttctgc cttaggttat gatgcgcctc	300
cttctgtgcy accaatgaga cgaattcaga acatttttaa aataatctaa gcatcattga	360
agcagtaaca caaaaaaaag gttcacttat ttctttttat tataacttac atcctttcaa	420
ataagtcttt gccctcatga agatttttta gaggaaagaa aggaataata gtattttcca	480
gttttgcttg acagttttcta aacaaaccaa aattatctca atpaaaggaa agatgtttct	540

This Page Blank (uspto)